

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 18:46:43 : Search time 1824.97 Seconds
(without alignments)
6834.203 million cell updates/sec

Title: US-09-508-832-9
Perfect score: 596
Sequence: 1 atggcaagcaactctctga.....ctgggtggtggaatgcattg 596

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_higo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
------------	-------	--------------------	-------	-------------

1	596	100.0	596	6	AX031287	Sequence
2	596	100.0	596	13	AX031313	Sequence
3	556	100.0	597	9	AF032457	Homo sapi
4	553	92.8	630	9	AB071197	Homo sapi
5	499.4	83.8	722	9	AB071195	Homo sapi
6	448.4	75.2	590	6	AX031283	Sequence
7	448.4	75.2	590	13	AX031309	Sequence
8	448.4	75.2	591	10	AF032459	Mus muscu
9	448.4	75.2	1137	10	BC019556	Mus muscu
10	443.6	74.4	591	10	AF065433	Rattus no
11	394.4	66.2	493	9	AB071198	Homo sapi
12	394.4	66.2	181497	9	AC096670	Homo sapi
13	294	49.3	416	6	AX031285	Sequence
14	294	49.3	416	13	AX031311	Sequence
15	294	49.3	417	9	AF032458	Homo sapi
16	207.6	34.8	422	6	AX031281	Sequence
17	207.6	34.8	422	13	AX031307	Sequence
18	207.6	34.8	423	10	AF032460	Mus muscu
19	204.4	34.3	423	10	AF136927	Rattus no
20	201.4	33.8	443	9	AB071199	Homo sapi
21	197.4	33.1	542	9	AB071196	Homo sapi
22	190.2	31.9	282	10	AF065431	Rattus no
23	130.2	21.8	232	9	AB071200	Homo sapi
24	127	21.3	332	6	AX031279	Sequence
25	127	21.3	332	13	AX031305	Sequence
26	127	21.3	333	10	AF032461	Mus muscu
27	123.8	20.8	333	10	AF065432	Rattus no
28	108.8	18.3	2050	9	AF455755	Homo sapi
29	108.8	18.3	150035	30	AC046192	Homo sapi
30	108.8	18.3	166638	2	AC108463	Homo sapi
31	99	16.6	150035	30	AC046192	Homo sapi
32	65.4	11.0	59827	2	AC104222	Mus muscu
33	43	7.2	125020	9	AF429315	Homo sapi
34	39.2	6.6	181221	2	AC099532	Bos tauru
35	39.2	6.6	299350	1	SM5591786	Sinorhizo
36	37.8	6.3	770	9	HSPAX75	H. sapiens
37	37.4	6.3	120488	9	AC073257	Homo sapi
38	37.4	6.3	138119	2	AC021655	Homo sapi
39	37.4	6.3	151094	2	AC091797	Felis cat
40	37.4	6.3	171758	2	AC091761	Felis cat
41	37.4	6.3	172560	2	AC091794	Felis cat
42	37.4	6.3	185735	9	AC078785	Homo sapi
43	37.4	6.3	194006	9	AC074044	Homo sapi
44	37.4	6.3	204143	2	AC096461	Rattus no
45	37.4	6.3	231952	2	AC024108	Homo sapi

ALIGNMENTS

RESULT 1
AX031287
LOCUS AX031287
DEFINITION Sequence 9 from Patent WO9914321.
ACCESSION AX031287
VERSION AX031287.1 GI:10278618
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

unidentified.
unclassified.

1 (bases 1 to 596)
O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S.,
Huang, D. C. and Strasser, A.
Novel therapeutic molecules
Patent: WO 9914321-A 9 25-MAR-1999;
INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HANSA (AU) ; REILLY
LORRANE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)

Location/Qualifiers
1..596
/organism="unidentified"
/db_xref="taxon:32644"
1..594

Linear PAT 20-SEP-2000
APP

/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09657.1"
/db_xref="GI:10278619"
/translation="MAKOPSDVSECDREGLOLAERPPOLRPGAPTSLOTPEQGNP
EGNHGGEDSCPHGSPGGLAPASGPGPATRSPLIFIMRSSLSSSSGYSFSDTD
RSPAPMSDCKSTQTPSPPCOAFNHYLSAMASHROAPADMRPEIWIJELRIQDEFN
ATYARVFLNNYQAEEDHPRVILLRLRYIVRLVRMH"
BASE COUNT 145 a 175 c 146 g 130 t
ORIGIN

Query Match 100.0%; Score 596; DB 6; Length 596;
Best Local Similarity 100.0%; Pred. No. 2.7e-149;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 atggcaagcaaccttcgatgaagttctgagttgacagagaaggtagacaattgcag 60
Db 1 ATGGCAAGCAACCTCTGATGTAAGTCTGAGTGTGACGAGAGGTAGACAATTGCAG 60
Oy 61 cctgcggagagcctccccagctcagacctggggccctacctccctacagacagacca 120
Db 61 CCTGCGGAGAGCCTCCCCAGCTCAGACCTGGGGCCCTACCTCCCTACAGACAGACCA 120
Oy 121 caaggtaactcctgaaggaatcagagagtgagggagacagctgccccacagcagccct 180
Db 121 CAAGGTAACTCTGAGGCAATCAGGAGGTGAAGGGGACAGCTGCCCCACGGCAGCCCT 180
Oy 181 caggcccgctggccccacctggccagccctggcccttttgcatacagatccccgcttttc 240
Db 181 CAGGGCCGCTGGCCCCACCTGCCAGCCCTGGCCCTTTTGTACACAGATCCCGGCTTTTC 240
Oy 241 atctttatgagaagatcctccctcgtctcgtatccctcagtggttatctctctttgac 300
Db 241 ATCTTTATGAGAAGATCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTCTTTTAC 300
Oy 301 acagacagagccagccacccatgagttgacaaatcaacaaacccacagctccct 360
Db 301 ACAGACAGAGCCAGCCACCATGATGTTGTGACAAATCAACANMACCCACCAAGTCTCT 360
Oy 361 tgcagggcttcaacacatactcagtgcaatggcttccatgagggcggtgaacctgca 420
Db 361 TGCAGGGCTTCAACACATATCTCAGTGCAATGGCTTCCATGAGGAGGCTGAACCTGCA 420
Oy 421 gatatgcgccagagatatgagtcgcccagagtgctggcggtatcgagagagtttaac 480
Db 421 GATATGCGCCAGAGATATGATGCGCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAA 480
Oy 481 gcttactatgcaagagaggtattttgaataattaccagagccgaagaccacacga 540
Db 481 GCTTACTATGCAAGAGAGGTATTTTGNATNTATCCAGCAGCCCAAGCCACCCACCA 540
Oy 541 atggttatctacgactgttacgttacattgtccgctggtggtggagaatgcatg 596
Db 541 ATGTTTATCTTACGACTGTTCAGTTACATTTGTCGGCTGCTGTTGGAGAAATGCATTG 596

RESULT 2
AX031313
LOCUS AX031313 596 bp DNA linear UNA 20-SEP-2000
DEFINITION Sequence 9 from Patent WO9914321.
ACCESSION AX031313
VERSION AX031313.1 GI:10278641
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 596)
O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S.,
Huang, D.C. and Strasser, A.
Novel therapeutic molecules
Patent: WO 9914321-A 25-MAR-1999;
JOURNAL INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HANSA (AU) ; REILLY

* LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
FEATURES
source
Location/Qualifiers
1..596
/organism="unidentified"
/db_xref="taxon:32644"
1..594
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09662.1"
/db_xref="GI:10278642"
/translation="MAKOPSDVSECDREGLOLAERPPOLRPGAPTSLOTPEQGNP
EGNHGGEDSCPHGSPGGLAPASGPGPATRSPLIFIMRSSLSSSSGYSFSDTD
RSPAPMSDCKSTQTPSPPCOAFNHYLSAMASHROAPADMRPEIWIJELRIQDEFN
ATYARVFLNNYQAEEDHPRVILLRLRYIVRLVRMH"
BASE COUNT 145 a 175 c 146 g 130 t
ORIGIN

Query Match 100.0%; Score 596; DB 13; Length 596;
Best Local Similarity 100.0%; Pred. No. 2.7e-149;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 atggcaagcaaccttcgatgaagttctgagttgacagagaaggtagacaattgcag 60
Db 1 ATGGCAAGCAACCTCTGATGTAAGTCTGAGTGTGACGAGAGGTAGACAATTGCAG 60
Oy 61 cctgcggagagcctccccagctcagacctggggccctacctccctacagacagacca 120
Db 61 CCTGCGGAGAGCCTCCCCAGCTCAGACCTGGGGCCCTACCTCCCTACAGACAGACCA 120
Oy 121 caaggtaactcctgaaggaatcagagagtgagggagacagctgccccacagcagccct 180
Db 121 CAAGGTAACTCTGAGGCAATCAGGAGGTGAAGGGGACAGCTGCCCCACGGCAGCCCT 180
Oy 181 caggcccgctggccccacctggccagccctggcccttttgcatacagatccccgcttttc 240
Db 181 CAGGGCCGCTGGCCCCACCTGCCAGCCCTGGCCCTTTTGTACACAGATCCCGGCTTTTC 240
Oy 241 atctttatgagaagatcctccctcgtctcgtatccctcagtggttatctctctttgac 300
Db 241 ATCTTTATGAGAAGATCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTCTTTTAC 300
Oy 301 acagacagagccagccacccatgagttgacaaatcaacaaacccacagctccct 360
Db 301 ACAGACAGAGCCAGCCACCATGATGTTGTGACAAATCAACANMACCCACCAAGTCTCT 360
Oy 361 tgcagggcttcaacacatactcagtgcaatggcttccatgagggcggtgaacctgca 420
Db 361 TGCAGGGCTTCAACACATATCTCAGTGCAATGGCTTCCATGAGGAGGCTGAACCTGCA 420
Oy 421 gatatgcgccagagatatgagtcgcccagagtgctggcggtatcgagagagtttaac 480
Db 421 GATATGCGCCAGAGATATGATGCGCCCAAGAGTTGCGGCGTATCGGAGACGAGTTTAA 480
Oy 481 gcttactatgcaagagaggtattttgaataattaccagagccgaagaccacacga 540
Db 481 GCTTACTATGCAAGAGAGGTATTTTGNATNTATCCAGCAGCCCAAGCCACCCACCA 540
Oy 541 atggttatctacgactgttacgttacattgtccgctggtggtggagaatgcatg 596
Db 541 ATGTTTATCTTACGACTGTTCAGTTACATTTGTCGGCTGCTGTTGGAGAAATGCATTG 596

RESULT 3
AF032457
LOCUS AF032457 597 bp mRNA linear PRI 19-FEB-1998
DEFINITION Homo sapiens BimEL mRNA, complete cds.
ACCESSION AF032457
VERSION AF032457.1 GI:2895495
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

LOCUS AX031285 416 bp DNA linear PAT 20-SEP-2000
 DEFINITION Sequence 7 from Patent WO9914321.
 ACCESSION AX031285
 VERSION AX031285.1 GI:10278616
 KEYWORDS
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 416)
 AUTHORS O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S., Huang, D.C. and Strasser, A.
 TITLE Novel therapeutic molecules
 JOURNAL Patent: WO 9914321-A 7 25-MAR-1999;
 INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
 FEATURES
 source /organism="unclassified"
 /db_xref="taxon:32644"
 1..>416
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC09656.1"
 /db_xref="GI:10278616"
 /translation="MAKOPSDVSSCDREGROLQPAERPPQLRPGAPTSLQTEPDQRS PAPMSCDKSTQTPSPPCOAFNHYLSAMSMRQAEADMRPEIMIAQLRRLRIGDEFNAY YARRVFLNYYQAEDHPRMVILRLRLRIVLVRMH"
 BASE COUNT 113 a 113 c 103 g 87 t
 ORIGIN
 Query Match 49.3%; Score 294; DB 6; Length 416;
 Best Local Similarity 100.0%; Pred. No. 2.9e-58;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 303 agacagagccagaccatgagttgtgacaaatcaacaaaccccaagtcctcttg 362
 DB 123 AGACAGAGCCAGCACCAGTGTGTGACAAATCAACAAACCCCAAGTCTCTTG 182
 QY 363 ccaggccttcacacatctatctcagtgcaatggtctccatgagcgagctgaacctgcaga 422
 DB 183 CCAGGCTTCAACCACTATCTCAGTGCATGCTTCCATGAGCGAGCTGACCTGCAGA 242
 QY 423 tatgcgccagagatagatgcgccaagatgcgctgctgagcgagctgaattcaacgc 482
 DB 243 TATGCGCCAGAGATATGATGTCGCCAAGAGTTCGGCGTATCGAGACGAGTTTACGC 302
 QY 483 ttactatgcaggagggtatttttgaataattaccagcagccgagacccacacgaat 542
 DB 303 TTACTATGCAAGGAGGTATTTTGAATAATTACCAAGCAGCCGAGACCCACCAAT 362
 QY 543 ggtattctacgactttacgttacattgtccgctggtgtgagaaatgcattg 596
 DB 363 GGTATCTTACGACTGTACGTTACATTTCCGCTGCTGTCGAGAAATGCATTG 416
 RESULT 14
 AX031311
 LOCUS AX031311 416 bp DNA linear UNA 20-SEP-2000
 DEFINITION Sequence 7 from Patent WO9914321.
 ACCESSION AX031311
 VERSION AX031311.1 GI:10278639
 KEYWORDS
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 416)
 AUTHORS O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S., Huang, D.C. and Strasser, A.
 TITLE Novel therapeutic molecules
 JOURNAL Patent: WO 9914321-A 25-MAR-1999;
 INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY

LOCUS AX031285 416 bp DNA linear PAT 20-SEP-2000
 DEFINITION Sequence 7 from Patent WO9914321.
 ACCESSION AX031285
 VERSION AX031285.1 GI:10278616
 KEYWORDS
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 416)
 AUTHORS O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S., Huang, D.C. and Strasser, A.
 TITLE Novel therapeutic molecules
 JOURNAL Patent: WO 9914321-A 7 25-MAR-1999;
 INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)
 FEATURES
 source /organism="unclassified"
 /db_xref="taxon:32644"
 1..>416
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC09661.1"
 /db_xref="GI:10278640"
 /translation="MAKOPSDVSSCDREGROLQPAERPPQLRPGAPTSLQTEPDQRS PAPMSCDKSTQTPSPPCOAFNHYLSAMSMRQAEADMRPEIMIAQLRRLRIGDEFNAY YARRVFLNYYQAEDHPRMVILRLRLRIVLVRMH"
 BASE COUNT 113 a 113 c 103 g 87 t
 ORIGIN
 Query Match 49.3%; Score 294; DB 13; Length 416;
 Best Local Similarity 100.0%; Pred. No. 2.9e-68;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 303 agacagagccagaccatgagttgtgacaaatcaacaaaccccaagtcctcttg 362
 DB 123 AGACAGAGCCAGCACCAGTGTGTGACAAATCAACAAACCCCAAGTCTCTTG 182
 QY 363 ccaggccttcacacatctatctcagtgcaatggtctccatgagcgagctgaacctgcaga 422
 DB 183 CCAGGCTTCAACCACTATCTCAGTGCATGCTTCCATGAGCGAGCTGACCTGCAGA 242
 QY 423 tatgcgccagagatagatgcgccaagatgcgctgctgagcgagctgaattcaacgc 482
 DB 243 TATGCGCCAGAGATATGATGTCGCCAAGAGTTCGGCGTATCGAGACGAGTTTACGC 302
 QY 483 ttactatgcaggagggtatttttgaataattaccagcagccgagacccacacgaat 542
 DB 303 TTACTATGCAAGGAGGTATTTTGAATAATTACCAAGCAGCCGAGACCCACCAAT 362
 QY 543 ggtattctacgactttacgttacattgtccgctggtgtgagaaatgcattg 596
 DB 363 GGTATCTTACGACTGTACGTTACATTTCCGCTGCTGTCGAGAAATGCATTG 416
 RESULT 15
 AF032458
 LOCUS AF032458 417 bp mRNA linear PRI 19-FEB-1998
 DEFINITION Homo sapiens BimL mRNA, complete cds.
 ACCESSION AF032458
 VERSION AF032458.1 GI:2895497
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 417)
 AUTHORS O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M., Cory, S. and Huang, D.C.
 TITLE Bim: a novel member of the Bcl-2 family that promotes apoptosis
 JOURNAL EMBO J. 17 (2), 384-395 (1998)
 MEDLINE 98094360
 PUBMED 9430630
 REFERENCE 2 (bases 1 to 417)
 AUTHORS O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M., Cory, S. and Huang, D.C.S.
 TITLE Direct Submission
 JOURNAL Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter & Eliza Hall Institute of Medical Research, PO Royal Melbourne Hospital, Parkville, Victoria 3050, Australia
 FEATURES
 source /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..417
 CDS

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 19:36:18 : Search time 210.72 Seconds
(without alignments)
4856.117 Million cell updates/sec

Title: US-09-508-832-9
Perfect score: 596
Sequence: 1 atggcaagcaactctctga.....ctgggtggaagatgcattg 596

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1004
Listing first 45 summaries

Database :			
N_Geneseq_032802.*			
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*		
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*		
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*		
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*		
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*		
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*		
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*		
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*		
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*		
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*		
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*		
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*		
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*		
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*		
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*		
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*		
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*		
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*		
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*		
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*		
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*		
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*		
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*		
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	596	100.0	596	20	Human Bcl-2 intera
2	448.4	75.2	590	20	Murine Bcl-2 inter
3	294	49.3	416	20	Human Bcl-2 intera
4	276	46.3	442	23	DNA encoding novel
5	207.6	34.8	422	20	Murine Bcl-2 inter
6	127	21.3	332	20	Murine Bcl-2 inter
7	71	11.9	371	22	Human polynucleoti
8	37	6.2	4047	21	Human Jurkat cell
9	37	6.2	4047	21	Human Jurkat cell

10	36.8	6.2	480	22	ABA42934	Human breast cell
11	36.8	6.2	480	22	ABA53355	Human foetal liver
12	36.8	6.2	480	22	ABA23135	Probe #1601 for ge
13	36.8	6.2	480	22	AAK01622	Human brain expres
14	36.8	6.2	480	22	AAK27064	Human bone marrow
15	36.8	6.2	480	22	AAI11662	Probe #1595 for ge
16	36.8	6.2	480	22	AAI32961	Probe #1647 used t
17	36.8	6.2	480	22	AAI01587	Probe #1578 used t
18	35.8	6.0	569	22	ABA63685	Human foetal liver
19	35.8	6.0	569	22	ABA30875	Probe #9341 for ge
20	35.8	6.0	569	22	AAK37923	Human brain expres
21	35.8	6.0	569	22	AAI18684	Human bone marrow
22	35.8	6.0	569	22	AAI43797	Probe #8617 for ge
23	35.8	6.0	569	22	AAI43797	Probe #12483 used
24	35.8	6.0	5352	22	AA513655	Genomic DNA sequen
25	35.6	6.0	16235	22	AAK86192	Human immune/haema
26	35.6	6.0	38186	20	AAZ32028	Human METH1 relate
27	35.6	6.0	38186	22	AAZ32028	AC004449 cDNA clon
28	35.4	5.9	1128	21	AAZ59247	Human NR8alpha/TPO
29	34.4	5.8	52616	22	AAK70459	Human immune/haema
30	34.4	5.8	52616	22	AAK78930	Human immune/haema
31	33.4	5.6	662	22	ABA19359	Human nervous syst
32	33.2	5.6	2567	22	AA501379	Human INTERCEPT
33	33.2	5.6	3770	22	AA501371	Human INTERCEPT
34	33	5.5	1464	16	AAQ76119	Human fetal cerebe
35	33	5.5	2913	16	AAQ76121	Human fetal cerebe
36	33	5.5	2923	16	AAQ76120	Human fetal cerebe
37	33	5.5	3183	16	AAQ76122	Human fetal brain
38	32.8	5.5	1942	22	AAI36808	Human musculoskele
39	32.6	5.5	786	23	AA574228	DNA encoding novel
40	32.6	5.5	814	22	AA544974	CDNA encoding nove
41	32.4	5.4	813	24	AA598127	Human DNA for pote
42	32.4	5.4	1806	23	ABL20075	Drosophila melanog
43	32.4	5.4	2238	23	AA565829	DNA encoding novel
44	32.4	5.4	3328	22	AAI05135	Human reproductive
45	32.4	5.4	6196	23	ABL20074	Drosophila melanog

ALIGNMENTS

RESULT	1
AAZ24997	
ID	AAZ24997 standard; cDNA; 596 BP.
XX	
AC	AAZ24997;
XX	
DT	05-JUL-1999 (first entry)
DE	
XX	Human Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX	
KW	Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
KW	cell cycle; human; cancer; autoimmune disease;
KW	degenerative disease; therapy; contraceptive; splice variant;
KW	isoform: 55.
XX	
OS	Homo sapiens.
XX	
PN	W09914321-Al.
XX	
PD	25-MAR-1999.
XX	
PF	17-SEP-1998; 98WO-AU00772.
XX	
PR	24-SEP-1997; 97AU-0009373.
PR	17-SEP-1997; 97AU-0009263.
XX	
PA	(HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX	
PI	Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
PI	Puthalakath H, Strasser A;
XX	
DR	WPI; 1999-244030/20.

DR P-PSDB: AAW98158.
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 PT treatment
 XX
 PS Claim 7: Page 101-102; 145pp; English.
 XX The present sequence encodes the extra long form (EL) of human Bim,
 CC or Bcl-2 interacting mediator of cell death (see AAW98158), a novel
 CC member of the Bcl-2 family that is capable of inducing cell death
 CC (apoptosis) and which acts as a 'death-ligand' for certain members
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
 CC only Bcl-2 homology region which it encompasses is BH3. It is the
 CC only BH3-only protein for which splice variants exist. These
 CC result in the expression of a variety of isoforms, i.e. Bim-S,
 CC Bim-L and Bim-EL. cDNAs encoding human Bim-L and Bim-EL (see
 CC AAW98158) were isolated from embryo and liver cDNA libraries using
 CC mouse b1m cDNA. Murine Bim-S, Bim-L and Bim-EL isoforms (see
 CC AAW98154-56) are also provided. The human Bim gene maps to
 CC chromosome 2 at bands 2q12-2q13. Binding the dynein light
 CC chain was shown to regulate the pro-apoptotic activity of Bim.
 CC Bim-S, the splice variant which does not bind to dynein light
 CC chain, is a much more potent killer than either Bim-L or Bim-EL.
 CC The invention provides variants (see AAW98159-68) of murine and human
 CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate
 CC with a dynein light chain. The identification of Bim permits the
 CC identification and rational design of a range of products for use
 CC in therapy, diagnosis, antibody generation and involving modulation
 CC of physiological cell death. These therapeutic molecules may act
 CC as either antagonists or agonists of Bim's function and will be
 CC useful in cancer, autoimmune or degenerative disease therapy.
 CC Increased Bim expression or Bim activity is useful, e.g. for
 CC treatment or prophylaxis in conditions such as cancer and deletion
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
 CC expression of Bim activity is useful in regulating inhibition or
 CC prevention of cell death or degeneration such as under cytotoxic
 CC conditions during e.g. gamma-irradiation and chemotherapy or during
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
 CC hypoxia, degenerative diseases or for prolonging the survival of
 CC cells being transplanted for treatment of disease. Since Bim is
 CC expressed in germ cells, modulating Bim expression or Bim activity
 CC is useful, e.g. as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 XX
 SQ Sequence 596 BP; 145 A; 175 C; 146 G; 130 T; 0 other;

Query Match 100.0%; Score 596; DB 20; Length 596;
 Best Local Similarity 100.0%; Pred. No. 4.3e-178;
 Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggcaagcaacctctgatgtgaagttctgagtgacgagaggtgagcaattgag 60
 Db 1 atggcaagcaacctctgatgtgaagttctgagtgacgagaggtgagcaattgag 60
 Qy 61 ctgcgagagagcctcccaagctcagacctgggcccctacctccctacagacagcc 120
 Db 61 ctgcgagagagcctcccaagctcagacctgggcccctacctccctacagacagcc 120
 Qy 121 caaggttaactgaaggcaatcacggagggtgaaggagagctgcctccacagagcc 180
 Db 121 caaggttaactgaaggcaatcacggagggtgaaggagagctgcctccacagagcc 180
 Qy 181 caggggccgctggcccccacccgagccctggccctttgtaccagatcccccgctttc 240
 Db 181 caggggccgctggcccccacccgagccctggccctttgtaccagatcccccgctttc 240
 Qy 241 attcttatgagaagatcctcctctctctctctctctctctctctctctctctct 300
 Db 241 attcttatgagaagatcctcctctctctctctctctctctctctctctctct 300
 Qy 301 acagacagagagccagcccatgagttgtacaaatcaacacaccccaagctcct 360
 Db 301 acagacagagagccagcccatgagttgtacaaatcaacacaccccaagctcct 360

Db 301 acagacagagagccagcccatgagttgtacaaatcaacacaccccaagctcct 360
 Qy 361 tgcagagccttcaaccactatctcagtgcaatgcttccatgagcagctgaacctga 420
 Db 361 tgcagagccttcaaccactatctcagtgcaatgcttccatgagcagctgaacctga 420
 Qy 421 gatatgcgcccagagatgatgagtcgcccagagttgcggtatcgagacagagtttaac 480
 Db 421 gatatgcgcccagagatgatgagtcgcccagagttgcggtatcgagacagagtttaac 480
 Qy 481 gcttactatgcagagaggggtatttttgaataattaccacagcagccagacccacagca 540
 Db 481 gcttactatgcagagaggggtatttttgaataattaccacagcagccagacccacagca 540
 Qy 541 atggttatcttaccagctgttaagttacatttcgctggtggtgaggaatgattg 596
 Db 541 atggttatcttaccagctgttaagttacatttcgctggtggtgaggaatgattg 596

RESULT 2
 AAX24995
 ID AAX24995 standard; cDNA; 590 BP.
 XX AC AAX24995;
 XX 05-JUL-1999 (first entry)
 XX Murine Bcl-2 interacting mediator of cell death Bim-EL cDNA.
 DE Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; mouse; cancer; autoimmune disease;
 KW degenerative disease; therapy; contraceptive; splice variant;
 KW isoform; ss.
 XX Mus musculus.
 OS
 XX WO9914321-A1.
 XX 25-MAR-1999.
 XX 17-SEP-1998; 98WO-AU00772.
 XX 24-SEP-1997; 97AU-0009373.
 XX 17-SEP-1997; 97AU-0009263.
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 XX WPI; 1999-244030/20.
 DR P-PSDB; AAW98156.
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 PT treatment
 XX Claim 3; Page 96-97; 145pp; English.
 XX The present sequence encodes the extra long form (EL) of murine Bim,
 CC or Bcl-2 interacting mediator of cell death (see AAW98156), a novel
 CC member of the Bcl-2 family that is capable of inducing cell death
 CC (apoptosis) and which acts as a 'death-ligand' for certain members
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
 CC only Bcl-2 homology region which it encompasses is BH3. It is the
 CC only BH3-only protein for which splice variants exist. These
 CC result in the expression of a variety of isoforms, i.e. Bim-S,
 CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim
 CC isoforms were obtained from a T lymphoma cDNA library using human
 CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to
 CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have
 CC also been identified (see AAW98157-58). Binding the dynein light
 CC chain was shown to regulate the pro-apoptotic activity of Bim.
 CC Bim-S, the splice variant which does not bind to dynein light

chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain. The identification of Bim permits the identification and rational design of a range of products for use in therapy, diagnosis, antibody generation and involving modulation of physiological cell death. These therapeutic molecules may act as either antagonists or agonists of Bim's function and will be useful in cancer, autoimmune or degenerative disease therapy. Increased Bim expression or Bim activity is useful, e.g. for treatment or prophylaxis in conditions such as cancer and deletion of autoreactive lymphocytes in autoimmune disease. Decreased Bim expression of Bim activity is useful in regulating inhibition or prevention of cell death or degeneration such as under cytotoxic conditions during e.g. gamma-irradiation and chemotherapy or during HIV/AIDS or other viral infections, ischemia, myocardial infarction, hypoxia, degenerative diseases or for prolonging the survival of cells being transplanted for treatment of disease. Since Bim is expressed in germ cells, modulating Bim expression or Bim activity is useful, e.g. as a contraceptive or method of sterilization by preventing generation of fertile sperm.

XX Sequence 590 BP; 137 A; 178 C; 150 G; 125 T; 0 Other;

Query Match 75.2%; Score 448.4; DB 20; Length 590;
Best Local Similarity 86.9%; Pred. NO. 1.9e-131;
Matches 523; Conservative 0; Mismatches 61; Indels 18; Gaps 2;

QY 1 atggcaagcaacctcttgatgaattctgagtgacccgaggaaggtagacaaatgcag 60
DB 1 atggcaagcaacctcttgatgaattctgagtgacccgaggaaggtgacaaatgcag 60
QY 61 cctgcagagagcctcccaagctcagacccgagccctaccctccctacagacagacca 120
DB 61 cctgcagagagcctcccaagctcagacccgagccctaccctccctacagacagacca 120
QY 121 caaggtatctcgaaggcaatccgaggtgaaggagagctgccccccagcagccct 180
DB 121 caaggtatctcgaaggc-----gaaggagagcctgccccccagcagccct 168
QY 181 caggccctcgcgcgcacccctgcagccctcttctgtacagatcccccgtcttc 240
DB 169 caggccctcgcgcgcacccctgcagccctcttctgtacagatcccccgtcttc 228
QY 241 atcttatgaagaatccctcctgctgtctgtatctcagtggtgtattctctttgac 300
DB 229 atcttctgagaagaattctctcgtcctcgtcctcagtggtgtattctctttgac 288
QY 301 acagacaggagccagcaccatagttgtgacaatacaacacacaccccaagctcct 360
DB 289 acagacaggagccgacccatagttgtgacaagtaacacacaccccaagctcct 348
QY 361 tgcagagcttcaaccactatctcagtgcaatgcttccatgagcagagct-----gaa 414
DB 349 tgcagagcttcaaccactatctcagtgcaatgcttccatgagcagagcttcaggaggaa 408
QY 415 cctcagagatgcgccagagatagtgatgcgcccaagagtgctgcggctatcgagagcag 474
DB 409 cctcagagatgcgccagagatagtgatgcgccagagagtgctgcggctatcgagagcag 468
QY 475 tttaacttactatgaaggaggtattttgataataatcacaagcagcagacacac 534
DB 469 ttcaacgaacttacaaggagggtgtttgcaaatgattaccgagagctgaagaccac 528
QY 535 ccagcaatggttatcttacaactgttacgtttacattgtccgcctgggtgagagaatgat 594
DB 529 cctcaaatggttatcttacaactgttacgtttacattgtccgcctgggtgagagaatgat 588
QY 595 tg 596
DB 589 tg 590

RESULT 3
ID AAX24996 standard; cDNA: 416 BP.
XX AAX24996;
XX AAX24996;
DT 05-JUL-1999 (first entry)
XX Human Bcl-2 interacting mediator of cell death Bim-L cDNA.
DE Bim-L: Bcl-2 interacting mediator of cell death; apoptosis;
KW cell cycle; human; cancer; autoimmune disease;
KW degenerative disease; therapy; contraceptive; splice variant;
KW isoform; ss.
XX Homo sapiens.
OS WO9914321-A1.
PN 25-MAR-1999.
PD 17-SEP-1998: 98WO-AU00772.
PF 24-SEP-1997: 97AU-0009373.
PR 17-SEP-1997: 97AU-0009263.
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
PI Puthalakath H, Strasser A;
DR WPI: 1999-244030/20.
OR P-PSDB; AAW98157.
XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment
XX Claim 7: Page 99-100; 145pp: English.
XX The present sequence encodes the long form (L) of human Bim, or Bcl-2 interacting mediator of cell death (see AAW98157), a novel member of the Bcl-2 family that is capable of inducing cell death (apoptosis) and which acts as a 'death-ligand' for certain members of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the only BH3-only protein region which it encompasses is BH3. It is the result in the expression of a variety of isoforms, i.e. Bim-S, Bim-L and Bim-EL. cDNAs encoding human Bim-L and Bim-EL (see AAW98157) were isolated from embryo and liver cDNA libraries using mouse Bim cDNA. Murine Bim-S, Bim-L and Bim-EL isoforms (see AAW98154-56) are also provided. The human Bim gene maps to chromosome 2 at bands 2q12-2q13. Binding the dynein light chain was shown to regulate the pro-apoptotic activity of Bim. Bim-S, the splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain. The identification of Bim permits the identification and rational design of a range of products for use in therapy, diagnosis, antibody generation and involving modulation of physiological cell death. These therapeutic molecules may act as either antagonists or agonists of Bim's function and will be useful in cancer, autoimmune or degenerative disease therapy. Increased Bim expression or Bim activity is useful, e.g. for treatment or prophylaxis in conditions such as cancer and deletion of autoreactive lymphocytes in autoimmune disease. Decreased Bim expression of Bim activity is useful in regulating inhibition or prevention of cell death or degeneration such as under cytotoxic conditions during e.g. gamma-irradiation and chemotherapy or during HIV/AIDS or other viral infections, ischemia, myocardial infarction, hypoxia, degenerative diseases or for prolonging the survival of cells being transplanted for treatment of disease. Since Bim is

CC expressed in germ cells, modulating Bim expression or Bim activity
 CC is useful, e.g. as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 XX
 SQ Sequence 416 BP; 113 A; 113 C; 103 G; 87 T; 0 other;

Query Match 49.3%; Score 294; DB 20; Length 416;
 Best Local Similarity 100.0%; Pred. No. 1e-82;
 Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 303 agacagagccagcaccatggtgtgacaaatcaacacacaccccaagctctcttg 362
 DB 123 agacagagccagcaccatggtgtgacaaatcaacacacaccccaagctctcttg 182
 OY 363 ccaggcttcaaccatctcagtgcaatgggttccatgagggcaggtgaacctgcaga 422
 DB 183 ccaggcttcaaccatctcagtgcaatgggttccatgagggcaggtgaacctgcaga 242
 OY 423 tatgcgcccagagatggtgacgcccagaggttgccggtatcgagacgagtttaacgc 482
 DB 243 tatgcgcccagagatggtgacgcccagaggttgccggtatcgagacgagtttaacgc 302
 OY 483 ttactatgaagaggttatcttgaaataattaccagcagcgaagaccacccacgaat 542
 DB 303 ttactatgaagaggttatcttgaaataattaccagcagcgaagaccacccacgaat 362
 OY 543 ggttatcttaacactgttacattacattgtccgctgggtggaatgcattg 596
 DB 363 ggttatcttaacactgttacattacattgtccgctgggtggaatgcattg 416

RESULT 4

AAS65350
 ID AAS65350 standard; cDNA; 442 BP.

XX AAS65350;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #1154.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; sa.
 XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG01163.

XX New isolated polynucleotide and encoded polypeptides, useful in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits and to assess
 XX biodiversity -

XX Claim 1; SEQ ID No 1154; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes.

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 442 BP; 94 A; 154 C; 105 G; 89 T; 0 other;

Query Match 46.3%; Score 276; DB 23; Length 442;

Best Local Similarity 100.0%; Pred. No. 5.1e-77;

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atggcaaaagcaacctctctgatgaagttctgagtgtagcagagaagtagacaattgcag 60

DB 167 atggcaaaagcaacctctctgatgaagttctgagtgtagcagagaagtagacaattgcag 226

OY 61 cctcgagagagcctccacagctcagacctggggccctacctccctcacagacagagcca 120

DB 227 cctcgagagagcctccacagctcagacctggggccctacctccctcacagacagagcca 286

OY 121 caaggttaactcgaaggaacacacagagtggaagggagcagctgccccacggcagccct 180

DB 287 caaggttaactcgaaggaacacacagagtggaagggagcagctgccccacggcagccct 346

OY 181 cagggcccgctggccccacctgcccagccctggcccttttctaccagatcccccgcttttc 240

DB 347 cagggcccgctggccccacctgcccagccctggcccttttctaccagatcccccgcttttc 406

OY 241 acctttatgagaagatccctccctgtgtctcgatcc 276

DB 407 acctttatgagaagatccctccctgtgtctcgatcc 442

RESULT 5

AAX24994
 ID AAX24994 standard; cDNA; 422 BP.

XX AAX24994;

XX 05-JUL-1999 (first entry)

XX Murine Bcl-2 interacting mediator of cell death Bim-L cDNA.

XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;

XX cell cycle; mouse; cancer; autoimmune disease;

XX degenerative disease; therapy; contraceptive; splice variant;

XX isoform; ss.

XX Mus musculus.

XX WO9914321-A1.

XX 25-MAR-1999.

XX 17-SEP-1998; 98WO-AU00772.

XX 24-SEP-1997; 97AU-0009373.

XX 17-SEP-1997; 97AU-0009263.

CC assess a disease condition associated with altered TRRE activity. The
 CC polypeptides, polynucleotides and antibodies can be used to decrease or
 CC increase signal transduction from a cytokine in a cell. The polypeptides,
 CC polynucleotides and antibodies may be used to treat heart failure,
 CC cachexia, inflammation, endotoxemic shock, arthritis, multiple sclerosis
 CC and sepsis, and cancer.

XX Sequence 4047 BP; 846 A; 1023 C; 1278 G; 900 T; 0 other;

Query Match 6.2%; Score 37; DB 21; Length 4047;
 Best Local Similarity 52.2%; Pred. No. 0.49;
 Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 65 cggagagagctcccccagctcagagctggggccctcctcctcagagagagcacaag 124
 DB 203 CGAGGCG 144
 QY 125 gtaatcctgaagcaatcagagctggagggagagctggcccccagcagcctcagg 184
 DB 143 GAGCTGCG 84
 QY 185 gccgcctggcccccagcagcctcctcctcctcctcctcctcctcctcctcctc 221
 DB 83 CCACTCTGCG 47

RESULT 9
 AAZ38863
 ID AAZ38863 standard; DNA: 4047 BP.
 XX AAZ38863;
 XX 17-FEB-2000 (first entry)
 XX Human Jurkat cell clone 2-9 AIM2 complementary nucleotide strand.
 XX Human; Jurkat cell; tumour necrosis factor receptor releasing enzyme;
 KW TRRE; cytokine; TNF; identification; cytostatic; anti-inflammatory;
 KW cardiant; immunomodulator; antiarthritis; antibacterial; cancer;
 KW heart failure; cachexia; inflammation; endotoxemic shock; arthritis;
 KW multiple sclerosis; sepsis; ds.
 XX Homo sapiens.

XX Location/Qualifiers
 FH KEY 378..1802
 FT CDS /*tag- a
 FT /*note- "longest open reading frame in the AIM2 clone"

XX WO9958559-A2.
 XX 18-NOV-1999.
 XX 14-MAY-1999; 99MO-US10793.
 XX 14-MAY-1998; 98US-0081385.
 XX (REGC) UNIV CALIFORNIA.
 XX Gatanaga T, Granger GA;
 XX WPI: 2000-039067/03.
 XX P-PSDB: AAY56509.

XX Tumor necrosis factor receptor releasing enzyme modulators and
 XX polynucleotides
 XX Example 5; Page 94-96; 106pp; English.
 XX The present invention describes isolated polynucleotides (A) comprising
 CC a sequence expressed at the mRNA level in Jurkat T cells and showing
 CC increased enzymatic activity for cleaving and releasing the tumour

CC necrosis factor (TNF) receptor in genetically modified COS-1 cells
 CC expressing the receptor. Methods from the present invention can be used
 CC to assess a disease condition associated with altered TRRE activity. The
 CC polypeptides, polynucleotides and antibodies can be used to decrease or
 CC increase signal transduction from a cytokine in a cell. The
 CC polypeptides, polynucleotides and antibodies may be used to treat heart
 CC failure, cachexia, inflammation, endotoxemic shock, arthritis, multiple
 CC sclerosis and sepsis, and cancer. The present sequence represents the
 CC complementary strand of the AIM2 clone which affects tumour necrosis
 CC factor receptor releasing enzyme (TRRE) activity, which encodes the
 XX longest open reading frame shown in AAY56509.

XX Sequence 4047 BP; 900 A; 1278 C; 1023 G; 846 T; 0 other;

Query Match 6.2%; Score 37; DB 21; Length 4047;
 Best Local Similarity 52.2%; Pred. No. 0.49;
 Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 65 cggagagagctcccccagctcagagctggggccctcctcctcagagagcacaag 124
 DB 3845 cggagc 3904
 QY 125 gtaatcctgaagcaatcagagctggagggagagctggcccccagcagcctcagg 184
 DB 3905 gagctggc 3964
 QY 185 gccgcctggcccccagcagcctcctcctcctcctcctcctcctcctcctcctc 221
 DB 3965 ccaacttgcccccagcagcctcctcctcctcctcctcctcctcctcctcctc 4001

RESULT 10
 ABA42934
 ID ABA42934 standard; DNA: 480 BP.
 XX ABA42934;
 XX 01-FEB-2002 (first entry)
 XX Human breast cell single exon nucleic acid probe #1629.
 XX Human; microarray; single exon probe; gene expression; breast;
 KW disease; cancer; ss.
 XX Homo sapiens.

XX WO200157271-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00662.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI: 2001-496933/54.
 XX New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes
 XX Claim 1; SEQ ID NO 1629; 327pp + sequence listing; English.

CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and Br 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequences. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 480 BP; 109 A; 162 C; 129 G; 80 T; 0 other;

Query Match 6.2%; Score 36.8; DB 22; Length 480;
Best Local Similarity 50.0%; Pred. No. 0.21; Mismatches 0; Gaps 0;
Matches 92; Conservative 0; Indels 0; Gaps 0;
Qy 65 cggagagggctcccccagctcagacctggggccctaccctccctacagagccacaag 124
Db 22 cgccgtgacttgaccaccagagcagtgcatgcagccagctggatgcaggtgcacaga 81
Qy 125 gtaactctgaaggcaatcagggagtgaggggagcagctgccccacagcgccctcagg 184
Db 82 tgatcccgccagctcttcacatcagggagcagcagctggggcccggtgcacagctcccg 141
Qy 185 gcccttgagccacagctgcagccctggcccttttctacagatccccgcttttcatct 244
Db 142 gctctcccccacacacacatcgtgggggaatttgctcccccaggccacagcttctcttc 201
Qy 245 ttat 248
Db 202 atat 205

RESULT 11
ABA53355
ID ABA53355 standard; DNA; 480 BP.
AC ABA53355;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #1660.
XX
KW Human: foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
PS Claim 1; SEQ ID NO 1660; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 480 BP; 109 A; 162 C; 129 G; 80 T; 0 other;

Query Match 6.2%; Score 36.8; DB 22; Length 480;
Best Local Similarity 50.0%; Pred. No. 0.21; Mismatches 0; Gaps 0;
Matches 92; Conservative 0; Indels 0; Gaps 0;
Qy 65 cggagagggctcccccagctcagacctggggccctaccctccctacagagccacaag 124
Db 22 cgccgtgacttgaccaccagagcagtgcatgcagccagctggatgcaggtgcacaga 81
Qy 125 gtaactctgaaggcaatcagggagtgaggggagcagctgccccacagcgccctcagg 184
Db 82 tgatcccgccagctcttcacatcagggagcagcagctggggcccggtgcacagctcccg 141
Qy 185 gcccttgagccacagctgcagccctggcccttttctacagatccccgcttttcatct 244
Db 142 gctctcccccacacacacatcgtgggggaatttgctcccccaggccacagcttctcttc 201
Qy 245 ttat 248
Db 202 atat 205

RESULT 12
ABA23135
ID ABA23135 standard; DNA; 480 BP.
XX
AC ABA23135;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #1601 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX


```
PR 21-SEP-2000: 2000US-0234687.
PR 27-SEP-2000: 2000US-0236359.
PR 04-OCT-2000: 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow.
XX Example 4: SEQ ID NO: 1621; 658pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX Sequence 480 BP; 109 A; 162 C; 129 G; 80 T; 0 other;
SQ
Query Match 5.2%; Score 36.8; DB 22; Length 480;
Best Local Similarity 50.0%; Pred. No. 0.21;
Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
Qy 65 cggagagggctcccagctcagacgtggggccctacccctccctacagacagagccacaag 124
Db 22 cgcgtggacttgaccacccagagcagtgcatgcccagctggtgagtgacaga 81
Qy 125 gtaactctgaagcaatcacggaggtgaaggagacagctgcccacggcagccctcagg 184
Db 82 tgatccgcagctcttcacctaggaggagcagcagctggggcccaagtgcacagctcccg 141
Qy 185 gccgcgtggcccacctgcagccctggcccttttctacacagatcccccgctttctatct 244
Db 142 gctgctcccccacccacacatcggtgggaattggtctccagggtcacagcttctcttc 201
Qy 245 ttat 248
Db 202 atat 205
RESULT 15
AA11662
ID AA11662 standard; DNA; 480 BP.
XX
AC AA11662:
XX
XX 12-OCT-2001 (first entry)
XX
XX Probe #1595 for gene expression analysis in human cervical cell sample.
XX Probe: human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; 86.
XX Homo sapiens.
XX
XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000670.
XX
XX 04-FEB-2000: 2000US-0180312.
XX 26-MAY-2000: 2000US-0207456.
XX 30-JUN-2000: 2000US-0608408.
XX 03-AUG-2000: 2000US-0632366.
XX 21-SEP-2000: 2000US-0234687.
XX 27-SEP-2000: 2000US-0236359.
```

```
PR 04-OCT-2000: 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells.
XX Claim 25; SEQ ID NO 1595; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 480 BP; 109 A; 162 C; 129 G; 80 T; 0 other;
SQ
Query Match 6.2%; Score 36.8; DB 22; Length 480;
Best Local Similarity 50.0%; Pred. No. 0.21;
Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
Qy 65 cggagagggctcccagctcagacgtggggccctacccctccctacagacagagccacaag 124
Db 22 cgcgtggacttgaccacccagagcagtgcatgcccagctggtgagtgacaga 81
Qy 125 gtaactctgaagcaatcacggaggtgaaggagacagctgcccacggcagccctcagg 184
Db 82 tgatccgcagctcttcacctaggaggagcagcagctggggcccaagtgcacagctcccg 141
Qy 185 gccgcgtggcccacctgcagccctggcccttttctacacagatcccccgctttctatct 244
Db 142 gctgctcccccacccacacatcggtgggaattggtctccagggtcacagcttctcttc 201
Qy 245 ttat 248
Db 202 atat 205
```

Search completed: August 7, 2002, 20:43:57
Job time: 4059 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 2002, 18:55:53 : Search time 46.29 Seconds
(without alignments)
3162.618 Million cell updates/sec

Title: US-09-508-832-9
Perfect score: 596
Sequence: 1 atggcaagcaacctctga.....ctggtgaggaatgcatg 596

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2.6/ptodata/1/ina/PTUS.COMB.seq.*
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	6.0	7218	1	US-08-232-463-14
2	33	5.5	1464	1	US-08-243-342-5
3	33	5.5	1464	1	US-08-477-407-5
4	33	5.5	1464	1	US-08-484-355-5
5	33	5.5	2913	1	US-08-243-342-7
6	33	5.5	2913	1	US-08-477-407-7
7	33	5.5	2913	1	US-08-484-355-7
8	33	5.5	2923	1	US-08-243-342-6
9	33	5.5	2923	1	US-08-477-407-6
10	33	5.5	2923	1	US-08-484-355-6
11	33	5.5	3183	1	US-08-243-342-8
12	33	5.5	3183	1	US-08-477-407-8
13	33	5.5	3183	1	US-08-484-355-8
14	32.2	5.4	3753	3	PCT-US95-02251-2
15	32.2	5.4	3759	3	US-08-479-7228-3
16	32.2	5.4	4314	1	US-08-199-780-2
17	32.2	5.4	4314	2	US-08-316-650-2
18	31.6	5.3	289	4	US-09-007-005-17
19	31.6	5.3	289	4	US-09-244-796-17
20	31.6	5.3	309	1	US-08-584-031-106
21	31.6	5.3	524	1	US-08-594-031-105
22	31.4	5.3	1509	1	US-08-216-276A-32
23	31.4	5.3	1622	1	US-08-216-276A-30
24	31.4	5.3	3230	1	US-08-216-276A-18
25	31.4	5.3	3230	1	US-08-219-2628-11
26	31.4	5.3	3230	3	US-09-031-655-11
27	31.4	5.3	3230	5	PCT-US91-03056-16

c 28	30.8	5.2	558	2	US-08-078-311-16	Sequence 16, Appl
c 29	30.8	5.2	558	2	US-08-460-402-16	Sequence 16, Appl
c 30	30.8	5.2	1907	1	US-08-184-327A-3	Sequence 3, Appl
c 31	30.8	5.2	1907	5	PCT-US95-00670-3	Sequence 3, Appl
c 32	30.8	5.2	2034	2	US-08-078-311-15	Sequence 15, Appl
c 33	30.8	5.2	2034	2	US-08-460-402-15	Sequence 15, Appl
c 34	30.6	5.1	1830	1	US-08-009-075-3	Sequence 3, Appl
c 35	30.2	5.1	2213	4	US-09-383-586-27	Sequence 27, Appl
c 36	30	5.0	576	4	US-09-385-982-23	Sequence 23, Appl
c 37	30	5.0	1941	3	US-09-082-737-1	Sequence 1, Appl
c 38	30	5.0	1987	2	US-08-990-379-1	Sequence 1, Appl
c 39	30	5.0	1993	2	US-08-990-379-2	Sequence 2, Appl
c 40	29.6	5.0	6226	1	US-08-542-363-1	Sequence 1, Appl
c 41	29.6	5.0	6226	4	US-09-100-089-1	Sequence 1, Appl
c 42	29.2	4.9	13987	2	US-08-804-227C-13	Sequence 13, Appl
c 43	29.2	4.9	44377	2	US-08-804-227C-7	Sequence 7, Appl
c 44	29.2	4.9	44377	2	US-08-804-198-1	Sequence 1, Appl
c 45	29.2	4.9	4403765	4	US-09-103-840A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEIFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTZgpt-Fls
: US-08-232-463-14

```

Query Match \      6.0%; Score 36; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. No. 0.12;
Matches 9; Conservative 138; Mismatches 93; Indels 0; Gaps 0;

Qy 162 ctccccccagcagccctcagggccgctggccacacctgcagccctggcccttttgc 221
Db 1212 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1271

Qy 222 tacagatccccgctttctatttatgagaagatccctcgtgctgcatactccag 281
Db 1272 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1331

Qy 282 tgggtattctcttttgacagacagaggagcaccatgagttgtgacaaacac 341
Db 1332 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1391

Qy 342 acaaaacccaagctcctctccagggcctcacaacctatctcagtgaatggctccat 401
Db 1392 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1451

RESULT 2
US-08-243-542-5
; Sequence 5, Application US/08243542
; Patent No. 5552526
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,542
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-45904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1464 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:

```

```

; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1464
; US-08-243-542-5

Query Match      5.5%; Score 33; DB 1; Length 1464;
Best Local Similarity 55.8%; Pred. No. 0.5;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 123 agttaacctgaagcaatcacgaggtgaaggagacagctgccccacagcagccctca 182
Db 164 ATGGGAACCTGACTTACATCGTGGAGCCCCCAAGAGGTGGCTGGACCTTGGGAGCCCTC 223

Qy 183 gggcccgctg9ggccacacctgcagcccttggccttggctaccagatccccgc 235
Db 224 AGGGACCCCTTCCACCTCATTACCGGACCCCTCTCCTCCAGATCCCTC 276

RESULT 3
US-08-477-407-5
; Sequence 5, Application US/08477407
; Patent No. 5631351
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,407
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,542
; FILING DATE: 13-MAY-1994
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-45904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1464 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE: EMI, MITSURU
ORGANISM: Homo sapiens
IMMEDIATE SOURCE: human fetal brain cdna library
LIBRARY: human fetal brain cdna library
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1464
US-08-477-407-5

Query Match
Best Local Similarity 55.8%; Score 33; DB 1; Length 1464;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 123 aggtatctgaaggaacatcagcggagtgaggggacagctgcccacagcagcctca 182
DB 164 ATGGGAACCTTACCTACATCGTGGAGCCCCCAAGAGGTGGCTGGAGCCCTC 223
OY 183 gggcccgctggccccacctgcagccctggcccttttgcctaccagatccccgc 235
DB 224 AGGAGACCCCTTCCCACTCATTTACCGGACCCCTCTCCTCCAGATCCCTC 276

RESULT 4
US-08-484-355-5
Sequence 5, Application US/08484355
Patent No. 5705341
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAs
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,355
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 1464 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE: EMI, MITSURU
ORGANISM: Homo sapiens
IMMEDIATE SOURCE: human fetal brain cdna library
LIBRARY: human fetal brain cdna library
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1464
US-08-484-355-5

Query Match
Best Local Similarity 55.8%; Score 33; DB 1; Length 1464;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 123 aggtatctgaaggaacatcagcggagtgaggggacagctgcccacagcagcctca 182
DB 164 ATGGGAACCTTACCTACATCGTGGAGCCCCCAAGAGGTGGCTGGAGCCCTC 223
OY 183 gggcccgctggccccacctgcagccctggcccttttgcctaccagatccccgc 235
DB 224 AGGAGACCCCTTCCCACTCATTTACCGGACCCCTCTCCTCCAGATCCCTC 276

RESULT 5
US-08-243-542-7
Sequence 7, Application US/08243542
Patent No. 5552526
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAs
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/243,542
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 7:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2913 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: 1..27
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 2038..2913
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..2037
;
; US-08-243-542-7
;
; Query Match 5.5%; Score 33; DB 1; Length 2913;
; Best Local Similarity 55.8%; Pred. No. 0.73;
; Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
;
; Qy 123 agttaatcctgaaggaatcaggaagtggaagggacagctccccccagcagccctca 182
; Db 212 ATGGGAACCTTGACTTACATCGTGGAGCCCAAGAGGTGGCTTGAGCCCTTGGGGAGCCCCCTC 271
;
; Qy 183 gggcccgctggccacacgtccctggcccttttgcaccagatccccgc 235
; Db 272 AGGAGCCCTTCCCCACCTCATTTACCGGACCCCTCTCTCCAGATCCCTC 324
;
; RESULT 6
; US-08-477-407-7
; Sequence 7, Application US/08477407
; Patent No. 5631351
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: EMI, MITSURU
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,407
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,542
; FILING DATE: 13-MAY-1994
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-49904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
```

```
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terrivence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2913 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: 1..27
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 2038..2913
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..2037
;
; US-08-477-407-7
;
; Query Match 5.5%; Score 33; DB 1; Length 2913;
; Best Local Similarity 55.8%; Pred. No. 0.73;
; Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
;
; Qy 123 agttaatcctgaaggaatcaggaagtggaagggacagctccccccagcagccctca 182
; Db 212 ATGGGAACCTTGACTTACATCGTGGAGCCCAAGAGGTGGCTTGAGCCCTTGGGGAGCCCCCTC 271
;
; Qy 183 gggcccgctggccacacgtccctggcccttttgcaccagatccccgc 235
; Db 272 AGGAGCCCTTCCCCACCTCATTTACCGGACCCCTCTCTCCAGATCCCTC 324
;
; RESULT 7
; US-08-484-355-7
; Sequence 7, Application US/08484355
; Patent No. 5705341
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: EMI, MITSURU
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,355
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,542
; FILING DATE: 13-MAY-1994
```



```

: APPLICATION NUMBER: JP 5-136602
: FILING DATE: 14 MAY 1993
: APPLICATION NUMBER: JP 5-257455
: FILING DATE: 22 SEPTEMBER 1993
: APPLICATION NUMBER: JP 6-49904
: FILING DATE: 23 FEBRUARY 1994
: APPLICATION NUMBER: JP 6-73328
: FILING DATE: 12 APRIL 1994
: APPLICATION NUMBER: JP 6-84470
: FILING DATE: 22 APRIL 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Terryence F. Chapman
: REGISTRATION NUMBER: 32 549
: REFERENCE/DOCKET NUMBER: Furuya Case 1313
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (616) 381-1156
: TELEFAX: (616) 381-5465
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2913 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: LIBRARY: human fetal brain cDNA library
: FEATURE:
: NAME/KEY: 5' UTR
: LOCATION: 1..27
: FEATURE:
: NAME/KEY: 3' UTR
: LOCATION: 2038..2913
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 28..2037
:
: US-08-484-355-7

Query Match 5.5%; Score 33; DB 1; Length 2913;
Best Local Similarity 55.8%; Pred. No. 0.73;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 123 agtgaatcctgaaggcaatcacgaggtgaaggagagagctgccccacggcagccctca 182
Db 212 ATGGGAACCTTGACTTACATCGTGAGCCCCCAAGAGTGCTGGACCTTGGGAGCCCCCTC 271
QY 183 gggcccgctggcccccacccctgccagccctgttgcacccagatccccgc 235
Db 272 AGGGACCCCTTCCGCCACCTCATTACCGGACCCCTCTCTCCACGATCCCCCTC 324

RESULT 8
US-08-243-542-6
: Sequence 6, Application US/08243542
: Patent No. 5552526
: GENERAL INFORMATION:
: APPLICANT: NAKAMURA, YUSUKE
: APPLICANT: : EMI, MITSURU
: TITLE OF INVENTION: MDC PROTEINS AND DNAs
: TITLE OF INVENTION: ENCODING THE SAME
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
: STREET: 2026 Rambling Road
: CITY: Kalamazoo
: STATE: Michigan
: COUNTRY: USA
: ZIP: 49008-1699
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
: COMPUTER: IBM PC/XT/AT Compatible

```

```

: OPERATING SYSTEM: MS-DOS 5.0
: SOFTWARE: WordPerfect 5.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/243,542
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 5-136602
: FILING DATE: 14 MAY 1993
: APPLICATION NUMBER: JP 5-257455
: FILING DATE: 22 SEPTEMBER 1993
: APPLICATION NUMBER: JP 6-49904
: FILING DATE: 23 FEBRUARY 1994
: APPLICATION NUMBER: JP 6-73328
: FILING DATE: 12 APRIL 1994
: APPLICATION NUMBER: JP 6-84470
: FILING DATE: 22 APRIL 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Terryence F. Chapman
: REGISTRATION NUMBER: 32 549
: REFERENCE/DOCKET NUMBER: Furuya Case 1313
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (616) 381-1156
: TELEFAX: (616) 381-5465
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2923 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: LIBRARY: human fetal brain cDNA library
: FEATURE:
: NAME/KEY: 5' UTR
: LOCATION: 1..27
: FEATURE:
: NAME/KEY: 3' UTR
: LOCATION: 1600..2923
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 28..1599
:
: US-08-243-542-6

Query Match 5.5%; Score 33; DB 1; Length 2923;
Best Local Similarity 55.8%; Pred. No. 0.73;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 123 agtgaatcctgaaggcaatcacgaggtgaaggagagagctgccccacggcagccctca 182
Db 212 ATCGGAACCTTGACTTACATCGTGAGCCCCCAAGAGTGCTGGACCTTGGGAGCCCCCTC 271
QY 183 gggcccgctggcccccacccctgccagccctgttgcacccagatccccgc 235
Db 272 AGGGACCCCTTCCGCCACCTCATTACCGGACCCCTCTCTCCACGATCCCCCTC 324

RESULT 9
US-08-477-407-6
: Sequence 6, Application US/08477407
: Patent No. 5631351
: GENERAL INFORMATION:
: APPLICANT: NAKAMURA, YUSUKE
: APPLICANT: : EMI, MITSURU
: TITLE OF INVENTION: MDC PROTEINS AND DNAs
: TITLE OF INVENTION: ENCODING THE SAME
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
: STREET: 2026 Rambling Road

```


OY 183 gggccgcgtggccacgtgcagccctggcccttttggctaccagatccccgc 235
DB 272 AGGAGCCCTTCCCGACCTCATTTACCGGACCCCTCTCTCCAGATCCCCCTC 324

RESULT 11

US-08-243-542-8
: Sequence 8, Application US/08243542
: Patent No. 5552526
: GENERAL INFORMATION:
: APPLICANT: NAKAMURA, YUSUKE
: APPLICANT: EMI, MITSURU
: TITLE OF INVENTION: MDC PROTEINS AND DNAs
: TITLE OF INVENTION: ENCODING THE SAME
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
: STREET: 2026 Rambling Road
: CITY: Kalamazoo
: STATE: Michigan
: COUNTRY: USA
: ZIP: 49008-1699
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
: COMPUTER: IBM PC/XT/AT Compatible
: OPERATING SYSTEM: MS-DOS 5.0
: SOFTWARE: WordPerfect 5.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/243,542
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 5-136602
: FILING DATE: 14 MAY 1993
: APPLICATION NUMBER: JP 5-257455
: FILING DATE: 22 SEPTEMBER 1993
: APPLICATION NUMBER: JP 6-49904
: FILING DATE: 23 FEBRUARY 1994
: APPLICATION NUMBER: JP 6-73328
: FILING DATE: 12 APRIL 1994
: APPLICATION NUMBER: JP 6-84470
: FILING DATE: 22 APRIL 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Terryence F. Chapman
: REGISTRATION NUMBER: 32 549
: REFERENCE/DOCKET NUMBER: Furuya Case 1313
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (616) 381-1156
: TELEFAX: (616) 381-5465
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3183 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: LIBRARY: human fetal brain cDNA library
: FEATURE:
: NAME/KEY: 3' UTR
: LOCATION: 2308..3183
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..2307
US-08-243-542-8

Query Match 5.5%; Score 33; DB 1; Length 3183;
Best Local Similarity 55.8%; Pred. No. 0.77;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 123 aggttaactctgaaggaatcaacgaggtgaaagggagacagctgccccacacggcagccctca 182
DB 482 ATGGGAACCTTGACTTACATCGTGGAGACCCCAAGAGAGGTGGCTGGACCTTGGGGAGCCCCCTC 541
OY 183 gggccgcgtggccacgtgcagccctggcccttttggctaccagatccccgc 235
DB 542 AGGAGCCCTTCCCGACCTCATTTACCGGACCCCTCTCTCCAGATCCCCCTC 594

RESULT 12

US-08-477-407-8
: Sequence 8, Application US/08477407
: Patent No. 5611351
: GENERAL INFORMATION:
: APPLICANT: NAKAMURA, YUSUKE
: APPLICANT: EMI, MITSURU
: TITLE OF INVENTION: MDC PROTEINS AND DNAs
: TITLE OF INVENTION: ENCODING THE SAME
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
: STREET: 2026 Rambling Road
: CITY: Kalamazoo
: STATE: Michigan
: COUNTRY: USA
: ZIP: 49008-1699
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
: COMPUTER: IBM PC/XT/AT Compatible
: OPERATING SYSTEM: MS-DOS 5.0
: SOFTWARE: WordPerfect 5.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/477,407
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/243,542
: FILING DATE: 13-MAY-1994
: APPLICATION NUMBER: JP 5-136602
: FILING DATE: 14 MAY 1993
: APPLICATION NUMBER: JP 5-257455
: FILING DATE: 22 SEPTEMBER 1993
: APPLICATION NUMBER: JP 6-49904
: FILING DATE: 23 FEBRUARY 1994
: APPLICATION NUMBER: JP 6-73328
: FILING DATE: 12 APRIL 1994
: APPLICATION NUMBER: JP 6-84470
: FILING DATE: 22 APRIL 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Terryence F. Chapman
: REGISTRATION NUMBER: 32 549
: REFERENCE/DOCKET NUMBER: Furuya Case 1313
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (616) 381-1156
: TELEFAX: (616) 381-5465
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3183 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: LIBRARY: human fetal brain cDNA library
: FEATURE:
: NAME/KEY: 3' UTR
: LOCATION: 2308..3183
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..2307
US-08-477-407-8

```
Query Match          5.5%  Score 33; DB 1; Length 3183;
Best Local Similarity 55.8%; Pred. No. 0.77; 50; Indels 0; Gaps 0;
Matches 63; Conservative 0; Mismatches 0; Gaps 0;

QY 123 aggtatctgaaggaatcacgaggtgaaggagacagctgccccacggcagccctca 182
Db 482 ATGGAACTTACTTACATCTGAGCCCAAGAGGTGCTGGACCTTGGGGAGCCCTC 541
QY 183 gggcccgctggccccacgtgcccccttttctaccagatccccgc 235
Db 542 AGGACCCCTTCCCACTTACGGACCCCTCTCTCTCCACATCCCTC 594

RESULT 13
US-08-484-355-8
; Sequence 8, Application US/08484355
; Patent No. 5705341
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; TITLE OF INVENTION: MDC PROTEINS AND DNAs
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,355
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,542
; FILING DATE: 13-MAY-1994
; APPLICATION NUMBER: JP 5-136602
; FILING DATE: 14 MAY 1993
; APPLICATION NUMBER: JP 5-257455
; FILING DATE: 22 SEPTEMBER 1993
; APPLICATION NUMBER: JP 6-49904
; FILING DATE: 23 FEBRUARY 1994
; APPLICATION NUMBER: JP 6-73328
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: JP 6-84470
; FILING DATE: 22 APRIL 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Terrence F. Chapman
; REGISTRATION NUMBER: 32 549
; REFERENCE/DOCKET NUMBER: Furuya Case 1313
; TELEPHONE: (616) 381-1156
; TELEFAX: (616) 381-5465
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3183 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: human fetal brain cDNA library
; FEATURE:
```

```
NAME/KEY: 3' UTR
LOCATION: 2308...3183
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2307
US-08-484-355-8

Query Match          5.5%  Score 33; DB 1; Length 3183;
Best Local Similarity 55.8%; Pred. No. 0.77;
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 123 aggtatctgaaggaatcacgaggtgaaggagacagctgccccacggcagccctca 182
Db 482 ATGGAACTTACTTACATCTGAGCCCAAGAGGTGCTGGACCTTGGGGAGCCCTC 541
QY 183 gggcccgctggccccacgtgcccccttttctaccagatccccgc 235
Db 542 AGGACCCCTTCCCACTTACGGACCCCTCTCTCTCCACATCCCTC 594

RESULT 14
PCT-US95-02251-2/C
; Sequence 2, Application PC/TUS9502251
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02251
; FILING DATE: CONCURRENTLY HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; CLASSIFICATION:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UNIC009P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3753 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3753
; PCT-US95-02251-2
```

Query Match 5.4%; Score 32.2; DB 5; Length 3753;
 Best Local Similarity 52.6%; Pred. No. 1.5;
 Matches 70; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
 OY 77 cccagctcagaccttggggccctacccctacagacagagccacaggtatctctgaag 136
 Db 200 CCTGTGTCAGCTGTCCTGACACTGGCCCTTCAGACAGGTCCGCTTCAGATCACAGGC 141
 OY 137 gcaatcacggaggtgaaggagagcagctgccccacggcagccctcagggcccgctggccc 196
 Db 140 GCAAAGACCACCTTGAAGCGTTGGGCCCGCCGCGCCCTGTGCCCCGCTGCC 81
 OY 197 caactgcccagccc 209
 Db 80 GGCGGCCCCACCC 68

RESULT 15
 US-08-479-722B-3/c
 ; Sequence 3, Application US/08479722B
 ; Patent No. 6074840
 ; GENERAL INFORMATION:
 ; APPLICANT: Bonadio, Jeffrey
 ; APPLICANT: Yin, Wushan
 ; TITLE OF INVENTION: LATENT TGF(BINDING PROTEIN (LTBP)
 ; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Williams, Morgan & Amerson
 ; STREET: 7676 Hillmont, Suite 250
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77040
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/479,722B
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US PCT/US95/02251
 ; FILING DATE: 21-FEB-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/316,650
 ; FILING DATE: 30-SEP-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/199,780
 ; FILING DATE: 18-FEB-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fussey, Shelley P. M.
 ; REGISTRATION NUMBER: 39,458
 ; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (713) 934-7000
 ; TELEFAX: (713) 934-7011
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3759 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..3759
 ; US-08-479-722B-3

Query Match 5.4%; Score 32.2; DB 3; Length 3759;

Best Local Similarity 52.6%; Pred. No. 1.5;
 Matches 70; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
 OY 77 cccagctcagaccttggggccctacccctacagacagagccacaggtatctctgaag 136
 Db 203 CCTGTGTCAGACAGCTGTCCTGACACTGGCCCTTCAGACAGGTCCGCTTCAGATCACAGGC 144
 OY 137 gcaatcacggaggtgaaggagagcagctgccccacggcagccctcagggcccgctggccc 196
 Db 143 GCAAAGACCACCTTGAAGCGTTGGGCCCGCCGCGCCCTGTGCCCCGCTGCC 84
 OY 197 caactgcccagccc 209
 Db 83 GGCGGCCCCACCC 71

Search completed: August 7, 2002, 20:40:18
 Job time: 6265 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic: - nucleic search, using sw model

Run on: August 7, 2002, 18:47:18 : Search time 1730.67 seconds
(without alignments)
4648.018 Million cell updates/sec

Title: US-09-508-832-9
Perfect score: 596
Sequence: 1 atggcaagaacactctctga.....ctgggtggagaatgcattg 596

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estba.*
2: em_estchun.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_gss.*
13: gb_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	430	72.1	1206	11	AK011490	Mus muscu
2	379.4	63.7	935	10	BG921698	BG921698 602825518
3	359.6	60.3	700	10	BI906766	BI906766 603064506
4	330	55.4	580	12	AZ706148	AZ706148 RPCT-23-2
5	329.4	55.3	645	9	BH651225	BH651225 BH651225
6	328.4	55.1	686	9	BH577135	BH577135 BH577135
7	315.4	52.9	452	10	BF021882	BF021882 UY59B09.Y
8	250.8	42.1	668	10	BG173095	BG173095 60233666
9	186.6	33.0	765	10	BI454637	BI454637 603170532
10	185.6	31.1	480	9	BH856627	BH856627 BH856627
11	185.2	31.1	684	9	BH653788	BH653788 BH653788
12	124	20.8	492	9	A1971169	A1971169 wr24h12.x
13	112.8	18.9	620	9	BH631162	BH631162 BH631162
14	112.8	18.9	912	10	BI555336	BI555336 603236095
15	99	16.6	135	10	BI497208	BI497208 df133e06
16	98	16.4	501	9	AA629308	AA629308 zu84g06.s
17	97	16.3	501	9	AA629050	AA629050 zu84a06.s

18	94.8	15.9	157	9	AW748960	AW748960 RC4-BT031
19	94.6	15.9	584	9	AW629314	AW629314 h156602.x
20	91.8	15.4	210	10	BF172831	BF172831 PCL5805.M
21	84.2	14.1	389	10	BF319454	BF319454 uy59b09.x
22	77.4	13.0	97	9	A1970428	A1970428 wr10d03.x
23	68.4	11.5	289	10	BG990772	BG990772 MRI-HT113
24	53.4	9.0	537	9	AF209718	AF209718 AF209718
25	41.4	6.9	1101	12	CNS00370	AL064465 Drosoph11
26	38.6	6.5	342	9	AW531398	AW531398 UI-R-BSO-
27	38.6	6.5	964	12	CNS003MG	AL065254 Drosoph11
28	38.4	6.4	925	12	CNS0091P	AL051013 Drosoph11
29	37.8	6.3	450	10	BES15545	BE515545 WHE0614.D
30	37.4	6.3	504	9	AV621963	AV621963 AV621963
31	37.4	6.3	639	10	BM003373	BM003373 103111080
32	37.4	6.3	651	10	BI718793	BI718793 1031031D1
33	37.4	6.3	662	10	BI724305	BI724305 1031069H1
34	37.4	6.3	685	10	BI717797	BI717797 1031022A1
35	37.4	6.3	1545	10	BM467095	BM467095 AGENCOURT
36	37.2	6.2	512	9	BB867390	BB867390 BB867390
37	37	6.2	643	10	BE295393	BE295393 601175911
38	37	6.2	717	10	BI261466	BI261466 602853519
39	37	6.2	894	10	BI114718	BI114718 602861747
40	36.8	6.2	298	10	BG275727	BG275727 601121880
41	36.8	6.2	773	10	BG421208	BG421208 602451695
42	36.4	6.1	587	9	AJ003454	AJ003454 AJ003454
43	36.4	6.1	925	12	CNS0091P	AL053013 Drosoph11
44	36.2	6.1	284	9	AI473314	AI473314 th83d09.x
45	36.2	6.1	414	10	BF850601	BF850601 PM1-END006

ALIGNMENTS

RESULT

AK011490

LOCUS

DEFINITION

AK011490

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

AK011490 1206 bp mRNA linear HTC 19-JAN-2002
Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2610020M23:BCL2-like 11 (apoptosis
facilitator), full insert sequence.

AK011490.1 GI:12847647

HTC: CAP trapper.

Mus musculus (strain:C57BL/6J) 10 days embryo cDNA to mRNA,

clone:lib:RIKEN full-length enriched mouse cDNA library

clone:2610020M23.

Mus musculus

Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;

Mammalia: Euthera: Rodentia: Sciurognathi: Muridae: Murinae; Mus.

1 (sites)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303:19-44 (1999)

99279253

10349636

2 (sites)

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

PUBMED

11042159

3 (sites)

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Hazada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Okada, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN Integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)


```

/db_xref="taxon:10090"
/clone_image=4954300
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/lab_stage="5 months"
/lab_host="DH10B"
/notes="organ: mammary; Vector: pCMV-SPORT6; Site_1: Self; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT      203 a      283 c      276 g      173 t
ORIGIN

Query Match      63.7%   Score 379.4; DB 10; Length 935;
Best Local Similarity 85.7%   Pred. No. 2.9e-97;
Matches 510; Conservative 0; Mismatches 97; Indels 24; Gaps 7;

QY 1 atggcaagcaacctcttgatgaagttctgagtgtagcaggaaggtagacaattgcag 60
    |||||
Db 208 ATGGCAAGCAACCTTCTGATGAAGTCTGAGTGTGACAGAGAAGGTGCACAAATTCGAG 267

QY 61 cctgcgagaggcctcccagctcagacctgaggcctggggccctaccctaccacagacagca 120
    |||||
Db 268 CTGCTGTAGAGGCGCTCCCCAGCTCAGGCGTGGGGCCCCCTACCTCCCTACAGACAGAACCG 327

QY 121 caaggtaatcctgaaggcaatcacggaggtgagggggacagctgccccacagcga-gccc 179
    |||||
Db 328 CAAGGTAAATCCGACGCGC-----GAAGGGGACCGCTGCCCCACCGCATGCC 375

QY 180 tcaggggcccgctggccccacctgcagccctggccc-ttttctaccagatccccgcttt 238
    |||||
Db 376 TCAGGGCCCGCTGGCCCCACCGCCCTGGCCCTTTTGTCTACCATGATCCCACTTT 435

QY 239 tcatttttagaagaatcctcctgtgtctgtatcgatcctcgaagtggtattctctttg 298
    |||||
Db 436 TCATCTTTGTGAGAAGATCTCTCTGTGTCGGCGTCTCCAGTGGGTATTCTCTTTTG 495

QY 299 acacagacgaggcccgagcacccatgagttgtgacaaatcaacacaaaccccaagctctc 358
    |||||
Db 496 ACNACAGAGGAGCGCCGGCACCACATGATGTGACAACTCAACACAAA-CCCAAGTCC 554

QY 359 ctgcacaggccttcaccactactcagtcgcaatgcttcacatgagcagagct- ----g 412
    |||||
Db 555 CTTGCCAGGCTTCAACCACTATCTCAGTGCATGCTTCCATAGCAGACTGCTCAGGAG 614

QY 413 aacctgcagata tgcgcccagagata tggatgcgcccaagagttgcggcgatcggagagc 472
    |||||
Db 615 AACCTGAAGATCTGCGCCCGGAGATACCGGATTCACAGGAGGCTGCGCGGATCGGAGAGC 674

QY 473 agtttaacgcttactatgcgaaggagggtatttttgaataattaccaagcagccgagacc 532
    |||||
Db 675 AGTTCAACGAAACTTACACAGGAGGGGTGTTGCAANTATTACCGCGAGCGCTCAAGA-C 733

QY 533 acccaaaatgggtattcttcaga--cgtgacattacatttgcgcgtcagtgtag 585
    |||||
Db 734 ACCCTCAATGGGTATTCTTTACAAAGGTTTACGCTTTATCTTCCTGCTGTGATGG 788

RESULT 3
BI906766
LOCUS      603064506F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5213713 5',
DEFINITION mRNA sequence.
ACCESSION BI906766
VERSION   BI906766
KEYWORDS  EST.
SOURCE    human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 700)

```

```

AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@email.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM1536 row: 1 column: 02
              High quality sequence stop: 696.
              Location/Qualifiers
                1..700
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5213713"
                  /clone_lib="NIH_MGC_118"
                  /tissue_type="leukocyte"
                  /lab_host="DH10B"
                  /note="vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRI
                    (destroyed); RNA source leukocytes from anonymous pool
                    non-activated adult donors. Library is oligo-dT primed
                    and directionally cloned (EcoRV site is destroyed upon
                    cloning). Average insert size 1.7 kb, insert size range
                    1.2-3.3 kb. Library is normalized and enriched for
                    full-length clones and was constructed by C. Gruber
                    (Invitrogen). Research Genetics tracking code 027. NO
                    this is a NIH_MGC Library."
              148 a 233 c 175g 144 t
BASE COUNT

```

Query Match	60.3%	Score	359.6	DB	10	Length	700
Best Local Similarity	96.2%	Prod. No.	1.1e-91				
Matches	433	Conservative	0	Mismatches	8	Indels	8
							Gaps
Oy	1	atgccaagcaac	cttctgagttaagt	-ctgagtgtgaccgagaaggtgagacaattgca	59		
Db	257	ATGCCAAACCA	CTTCGTGATGAAGTTCCTGTAGTGTGACCGAAGAGGTAGACAATTGCA	316			
Oy	60	gcttcgagagagc	ctcccgctcagacactgagacctggggcccttaacctcaacagacagagacc	119			
Db	317	GCTCGGAGAGG	CTCTCCAGCTCAGACCTCGGGGCCCTTACCTCCCTACAGACGAGGCC	376			
Oy	120	acaaggttaat	ctctgaaggcaatcacggaggtagaaggggacagctgcgcccaacggcagacc	179			
Db	377	ACAAGGTAAT	TCCTGAAGGCAATACGGAGGTGAAGGGGACAGCTGCGCCGACGGAGCGCTC	436			
Oy	180	tcaggccgcgc	gtggcccaactgcagccctggccctttctgctaccagatccccctttt	239			
Db	437	---AGGCGCG	GTGGGCCCACTGCCAGCCCTGG-CTTTTGTCTACCCAGATCCCGCGTTT	492			
Oy	240	catctttatga	agaagatcctccctgctgtctcgaatccctcagtgaggatattctctttga	299			
Db	493	CATCTTTATGA	AAAGATCTCTCTGCTGTCTGATCTCCAGTGGGTA-TTCTCTTTTGA	551			
Oy	300	cacagacaggg	accagcacccatgagttgacaaatacaacaacccaagctctcc	359			
Db	552	CACAGACAGG	AGCCCAGACAGCATGAGTTGTGACAAATCAACAACGCCAAGCTCTCC	611			
Oy	360	tgcacagcctc	acaccactatctcagtgcaatggcttccatgagcagctgaacctgc	419			
Db	612	TTGCCAGGCT	TCAA-CACTATCTCAGTGCANTGGCTTCCATGAGCAGGGCTGAACCTGG	670			
Oy	420	agatatgcgc	acagagatatggatgccc	449			
Db	671	AGATATGCC	-CCAGAGATATGGTCGCCCA	699			

RESULT 4

AZ706148
LOCUS
DEFINITION
AZ706148 580 bp DNA linear GSS 24-JAN-2001
RPIC1-23-227P3.TV RPIC1-23 Mus musculus genomic clone
DNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 580)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
Mouse BAC End Sequences from Library RPIC1-23
Unpublished (1999)
Other_GSSs: RPIC1-23-227P3.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPIC1-23. For BAC
library availability, please contact Pieter de Jong
(pdejong@tigr.org). Clones may be purchased from BACPAC
Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end
page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 227 row: P column: 3
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..580
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPIC1-23-227P3"
/clone_lib="RPIC1-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site:1:
EcoRI; Site:2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRII. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 138 a 162 c 138 g 142 t
ORIGIN

FEATURES

source

Query Match 55.4%; Score 330; DB 12; Length 580;
Best Local Similarity 91.9%; Pred. No. 2.9e-83;
Matches 364; Conservative 0; Mismatches 20; Indels 12; Gaps 1;
Qy 1 atggcaagcaacctctgagtgagttgagtgagcaggaaggtagacaattcag 60
Db 90 ATGCCCAACACACCTCTGATGTAGTCTGAGTGTGACAGAGAGAGGTGGACATTCGAG 149
Qy 61 cctgcggagagggcctccccagctcagacctg99ggccctacctctctacagagagccca 120
Db 150 CTTGCTGACAGAGCCCTCCCGACCTCAGGCGCTGGGGCCCTACCTCCCTACAGACAGAACCG 209
Qy 121 caaggttaatcctgagagcaatcacgaggggagcagctgccccccagcagcgcct 180
Db 210 CAAGGTAATCCCGACGCG-----GAAGGGGACCGCTGCGCCCGACGCGCGCT 257
Qy 181 caggcccgctgagccacctgagcagcctggccctttgtaccagatccccgcttttc 240
Db 258 CAGGCCCCCTGCCCCCAGCCGCGCCCTGGGCCCTTTTGTACACAGATCCCCACTTTTC 317

241 atctttatgagagatcctccctgctgctcgtatcgtatcctccagtggttctcttttgcac 300
Db 318 ATCTTTGTCAGAGAATCTTCTCTGCTGTCGCGGCTCTCCAGTGGGTATTCTCTTTTCAC 377
Qy 301 acagacagagagccagcaccatagttgtacacaaatcaacacaaaccccaagctcct 360
Db 378 ACAGACAGGAGCGCCGACCCCATGAGTTGTGACANGTCAACACCAACCCCAAGTCTCTCT 437
Qy 361 tgcagagcttcaacacactatctcagtcgaatggct 396
Db 438 TGCAGGCGCTTCAACCACTATCTCAGTGCATATGGGT 473
RESULT 5
BB651225 645 bp mRNA linear EST 26-OCT-2001
LOCUS
DEFINITION
BB651225 RIKEN full-length enriched, 0 day neonate cerebellum Mus
musculus cDNA clone C230089E03 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 645)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.
M., Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sesaki
D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Atakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10): 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Wadhvani, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11): 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2): 281-289 (2001)
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.
Location/Qualifiers
1..645
/organism="Mus musculus"
/db_xref="taxon:10090"

FEATURES

source

Oy 596 q 596
Db 594 g 594

RESULT 10

BB856627
LOCUS
DEFINITION
BB856627 RIKEN full-length enriched, B16 F10Y cells mRNA linear EST 26-NOV-2001
CDNA clone G370035M16 5', mRNA sequence.

ACCESSION

BB856627

VERSION

BB856627.1

KEYWORDS

EST

SOURCE

house mouse

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 480)

AUTHORS

Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,

Akimura, T., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii

, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,

Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,

Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa

, A., Takahashi, F., Takai-Akanishi, S., Tanaka, Y., Tomaru, A., Toya, T.,

Watanabe, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.

2001)

JOURNAL

unpublished (2001)

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, K., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura

, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and

Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara

, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

FEATURES

Location/Qualifiers

1..480

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="G370035M16"

/clone_lib="RIKEN full-length enriched, B16 F10Y cells"

/cell_type="B16 F10Y cells"

88 a 175 c 121 g 96 t

BASE COUNT

ORIGIN

Query Match 31.1% Score 185.6; DB 9; Length 480;

Best Local Similarity 87.5%; Pred. No. 4e-42;

Matches 231; Conservative 0; Mismatches 19; Indels 14; Gaps 2;

Oy 1 atggcaaaacacacccctctgattgaagttctgagtgtagccgagaagtagacaattgag 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 230 ATGCCAAGCAACCTTCGTGTAGTTCCTGAGTGTGAGAGAGAGAGAGAGAGAGAG 289
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Oy 61 cctg-cyggagggccctccacagctcagacctggggggccctacccctacacagagacc 119
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 290 CTTGCTAGAGAGCCCTCCAGCTCAGGCTGGGGCCCTACCTCCCTACAGACAGA--- 346
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Oy 120 acaaggaatacctcgaagcaatcacggaggtgaagggagacgtgcccccacggcagcc 179
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 347 -----ACGAAAGGTAAATCCGACGGGGAAGGGGACCGCTGCCCCACGGCAGCC 396
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Oy 180 tcaggggcccgctggcccaactgcccagccctggccctttgttaacagatccccctttt 239
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 397 TCAGGGCCCGCTGGGCCCCACCGGGCCAGCCCTGGCCCTTTTCTTACCAGATGCCCACTTTT 456
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Oy 240 catctttatgagaagatctctcct 263
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 457 CATCTTTGTGAGAAGATCTTCTCT 480
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 11

BB653788

LOCUS

BB653788 RIKEN full-length enriched, adult male liver tumor Mus

musculus cDNA clone C730026B05 5', mRNA sequence.

ACCESSION

BB653788

VERSION

BB653788.1

KEYWORDS

EST

SOURCE

house mouse

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 684)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.

, Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda

, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,

Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki

, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,

Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,

Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

JOURNAL

unpublished (2001)

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, K., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura

, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and

Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara

, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES
source
1. .684
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="C730026B05"
/clone_lib="RIKEN full-length enriched, adult male liver
tumor"
/sex="male"
/tissue_type="liver tumor"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: Sali; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCCGACCTGAGTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATCTCGAGTTAAATTAATTCCTCCCTCCCTCTG 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pluescript KS(+) after bulk excision from Lambda FLC I.
Tissue was provided by William A. Held, Roswell Park
Cancer Institute, Department of Molecular and Cellular
Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose
assistance we gratefully acknowledge."
BASE COUNT 162 a 208 c 174 g 140 t
ORIGIN
Query Match 31.1%; Score 185.2; DB 9; Length 684;
Best Local Similarity 83.0%; Pred. No. 6.1e-42;
Matches 249; Conservative 0; Mismatches 43; Indels 8; Gaps 3;
Oy 303 agacagagccagcaccatgattgtgacaaatcaacacaccccccctctctg 362
Db 337 AGACAGAGCCCGCACCCTGAGTTGTGACAGTCAACACACCCCAAGTCTCTTG 396
Oy 363 ccagggcttcaaccactatctcagtgcaatggcttccatgagcaggct-----gaacc 416
Db 397 CCAGGCTTCAACCACTATCTCAGTGCATGGCTTCCATACGACAGTCTCAGGAGAAC 456
Oy 417 tgcagatagcgcagagatagtgatcgcccaagaggttgccggtatcgagagaggtt 476
Db 457 TGAAGATCTCGCGCCGGAGATACGAGTTGCACAGGAGCTCGCGGATCGGAGACGATT 516
Oy 477 taacgttactatgcagagaggtatttttgaataataacacagcagcgaagacc 536
Db 517 CAACCAAACTTACAAAGAGAGGCTTTTGCAGAAATGATTACCGGAGGTGAGAC-CA 575
Oy 537 acgaatggtttatctacgactattacgttacattgtccgcctggtgaggaatgcattg 596
Db 576 TCAATGGTTATCTTAC-ACTGTTACGCTTTATCTTCCGTTGATGGAGAGGCATTG 634
RESULT 12
LOCUS AI971169/c 492 bp mRNA linear EST 08-MAR-2000
DEFINITION w24h12.x1 NCI_CGAP_P28 Homo sapiens cDNA clone IMAGE:248679 3'
similar to TR:043522 O43522 B1ML. [1] ;, mRNA sequence.
ACCESSION AI971169
VERSION AI971169.1 GI:5767995
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 492)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
Clone sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/brp/image/image.html
Insert Length: 712 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 450.
Location/Qualifiers
1. .492
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:248679"
/clone_lib="NCI_CGAP_P28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: p773D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_P28 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
985608-986759, 110192-110199, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 119 a 107 c 130 g 134 t
ORIGIN
Query Match 20.8%; Score 124; DB 9; Length 492;
Best Local Similarity 100.0%; Pred. No. 1.5e-24;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 atggcaagcaacctctctgattgattgagtgaccgagaggtagacaattgcag 60
Db 425 ATGGCAAGCAACCTCTGTATTAAGTTCTGAGTGTGACCGAGAGGTAGCAATTCGAC 366
Oy 61 cctgcgagagagcctcccaagctcagacacgtggggccctacctccctcacagagccca 120
Db 365 CCGCGGAGAGGCGCTCCCGAGCTCAGACCTGGGGCCCTACCTCCCTACACAGACGCCA 306
Oy 121 caag 124
Db 305 CAAG 302
RESULT 13
LOCUS BB631162 620 bp mRNA linear EST 26-OCT-2001
DEFINITION BB631162 RIKEN full-length enriched, 16 days neonate thymus Mus
musculus cDNA clone A130059E14 5', mRNA sequence.
ACCESSION BB631162
VERSION BB631162.1 GI:16467982
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 620)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanaoka, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Onda, K., Okazaki, Y., Okido, F., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished (2001)

TITLE
COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suchoiro-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 URL: <http://genome-gsc.riken.go.jp/>

1 (bases 1 to 620)
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 wadl, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Tamana, I., Alzawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome.* 12, 673-677 (2001)
 Please visit our web site (<http://genome-gsc.riken.go.jp>) for further details.

e mouse tissues.

Location/Qualifiers
 1. 620
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="RIKEN full-length enriched, 16 days neonate thymus"
 /tissue_type="thymus"
 /dev_stage="16 days neonate"
 /lab_host="DH109"
 /note="Site 1: Salt; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGAGAGAGAGATCCAGATCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot - 10.0 and subtraction to Rot - 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTGAGTTAATTAATATCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified phagescript KS(+) after bulk excision from Lambda

FEATURES
SOURCE

Location/Qualifiers
 1. 620
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="RIKEN full-length enriched, 16 days neonate thymus"
 /tissue_type="thymus"
 /dev_stage="16 days neonate"
 /lab_host="DH109"
 /note="Site 1: Salt; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGAGAGAGAGATCCAGATCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot - 10.0 and subtraction to Rot - 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTGAGTTAATTAATATCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified phagescript KS(+) after bulk excision from Lambda

Query Match 18.9% Score 112.8; DB 9; Length 620;
 Best Local Similarity 94.4% Pred. No. 2.6e-21;
 Matches 117; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 1 atggcgaagcaacctctctgattgagttctgagtgaccgagagagtgacaaattgcag 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 220 ATGCCCAAGCAACCTTCTGATCTAAGTTCTGAGTGTGACAGAGAGGTGGCAATTCGAC 279
 Oy 61 cctgcgagcggcctcccgagctcagcctggggccctccctccctccctccagcagcca 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Ub 280 CCTGCTGAGAGAGGCTCCCGAGCTCAGGCTGAGGCTGGGGCCCCCTACCTCCCTACAGACAGAACCG 339
 Oy 121 caag 124
 |||||
 Db 340 CAAG 343

RESULT 14

BI555336
 LOCUS 603236095F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5289406 5',
 DEFINITION mRNA sequence.

ACCESSION BI555336
 VERSION BI555336.1 GI:15442650
 KEYWORDS EST.
 SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC <http://mgi.nci.nih.gov/>.

REFERENCE

1 (bases 1 to 912)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@p51mail.nih.gov

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1731 row: m column: 23
 High quality sequence stop: 747.
 Location/Qualifiers
 1. 912

FEATURES

source
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5289406"
 /clone_lib="NIH_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /lab_host="DH109"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert 2 kb. Library constructed by Life Technologies, catalog #12017-018. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., *Nature Genetics* 22, 37-43 (1999). Note: this is a NCI_CGAP Library."
 BASE COUNT 216 a 242 c 267 g 187 t
 ORIGIN

Query Match 18.9% Score 112.8; DB 10; Length 912;
 Best Local Similarity 94.4% Pred. No. 3.1e-21;
 Matches 117; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 1 atggcgaagcaacctctctgattgagttctgagtgaccgagagagtgacaaattgcag 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 261 ATGCCCAAGCAACCTTCTGATCTAAGTTCTGAGTGTGACAGAGAGGTGGCAATTCGAC 320

BASE COUNT
ORIGIN

OM of: US-09-508-832-10 to: GenEmbl.* out_format : pfs

Date: Aug 7, 2002 10:35 PM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09508832/runat_06082002.103611.2778/app_query.fasta_1.259
-DB=GenEmbl -QEMT=Eastap -SUFFIX=ase -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPPE=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=0.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -NATRIK=BIOSUM62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09508832.ecgnl_1.3954
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-508-832-10
Query length: 198
Database: GenEmbl.*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 1829.680000

score_list:

Sequence	Strd	Orig	zscore	Escore	Len	Documentation
gb_pat:AX031287	+	1065.00	1039.47	1.6e-49	596	AX031287 Sequence 9 from Patent
gb_un:AX031313	+	1065.00	1039.47	1.6e-49	596	AX031313 Sequence 9 from Patent
gb_pat:AF032457	+	1065.00	1039.46	1.6e-49	597	AF032457 Homo sapiens BimEL mRNA
gb_pat:AF071197	+	1040.50	1024.15	1.1e-48	630	AF071197 Homo sapiens mRNA for
gb_pat:AF071195	+	1010.50	985.61	1.1e-46	722	AF071195 Homo sapiens mRNA for
gb_pat:AX031283	+	908.00	888.51	1.0e-41	590	AX031283 Sequence 5 from Patent
gb_un:AX031309	+	908.00	888.51	1.0e-41	590	AX031309 Sequence 5 from Patent
gb_un:AF032459	+	908.00	888.49	1.0e-41	591	AF032459 Mus musculus BimEL mRNA
gb_ro:BC019556	+	908.00	883.63	7.4e-41	1137	BC019556 Mus musculus, BCL2-1
gb_ro:AF065433	+	893.00	874.06	2.5e-40	591	AF065433 Rattus norvegicus Bcl-
gb_pat:AF071198	+	873.50	822.02	2.0e-37	493	AF071198 Homo sapiens mRNA for
gb_pat:AC096670	+	713.00	658.32	2.6e-28	181497	AC096670 Homo sapiens chromo
gb_un:AX031285	+	696.00	687.16	6.5e-30	416	AX031285 Sequence 7 from Patent
gb_pat:AX031311	+	696.00	687.16	6.5e-30	416	AX031311 Sequence 7 from Patent
gb_pat:AF032458	+	641.50	632.76	7.0e-27	542	AF032458 Homo sapiens BimL mRNA
gb_pat:AF071196	+	589.00	584.11	3.6e-24	422	AF071196 Homo sapiens mRNA for
gb_un:AX031281	+	589.00	584.11	3.6e-24	422	AX031281 Sequence 3 from Patent
gb_un:AF032460	+	589.00	584.10	3.6e-24	423	AF032460 Mus musculus BimL mRNA
gb_ro:AF032460	+	574.00	569.67	2.3e-23	423	AF032460 Rattus norvegicus Bcl-
gb_ro:AF032460	+	574.00	569.67	2.3e-23	423	AF032460 Rattus norvegicus Bcl-
gb_pat:AF071199	+	535.00	531.80	2.9e-21	443	AF071199 Homo sapiens mRNA for
gb_un:AX031305	+	406.00	409.85	1.8e-14	332	AX031305 Sequence 1 from Patent
gb_un:AX031279	+	406.00	409.85	1.8e-14	332	AX031279 Sequence 1 from Patent
gb_un:AX031261	+	388.00	392.51	1.7e-13	333	AX031261 Mus musculus BimS mRNA
gb_un:AF065432	+	388.00	388.93	2.7e-13	282	AF065432 Rattus norvegicus Bcl-
gb_un:AF065431	+	285.50	306.50	1.0e-08	223	AF065431 Rattus norvegicus Bcl-
gb_pat:AF071200	+	285.50	306.50	1.0e-08	223	AF071200 Homo sapiens mRNA for
gb_pat:AF071200	+	285.50	306.50	1.0e-08	223	AF071200 Homo sapiens mRNA for
em_hcg_hum:AC046192	-	192.00	158.53	2.01	150035	AC046192 Homo sapiens chromo
em_hcg_hum:AC046192	-	192.00	158.53	2.01	150035	AC046192 Homo sapiens chromo
em_hcg_hum:AC046192	-	170.00	137.36	27.44	150035	AC046192 Homo sapiens chromo
gb_hcg:AC104222	-	128.00	103.79	2.0e+03	59827	AC104222 Mus musculus clone h
gb_pat:AC004907	-	125.50	114.27	530.19	10570	AC004907 Pseudomonas aerugin
gb_pat:AC004971	-	125.50	93.13	8.0e+03	123253	AC004971 Homo sapiens PAC cl
gb_pat:HS407A10	-	121.50	101.78	2.6e+03	33804	HS407A10 Human DNA sequence fr
gb_hcg:OSJN01012	-	121.50	89.94	1.2e+04	166307	OSJN01012 Oryza sativa chromo
gb_hcg:AE006464	-	121.50	86.73	1.8e+04	256073	AE006464 Homo sapiens 16p13
gb_hcg:AC0023841	-	121.00	89.69	1.2e+04	161297	AC0023841 Homo sapiens chrom
gb_pat:AC005095	-	120.50	91.41	1.0e+04	119937	AC005095 Homo sapiens BAC cl
gb_hcg:AC095292	-	120.50	90.91	1.1e+04	128280	AC095292 Rattus norvegicus d

gb_pat:U82668 - 119.50 92.99 8.1e+03 85132 ! U82668 Homo sapiens shox g
gb_hcg:AC008749.2 - 119.50 91.09 1.0e+04 110000 ! Continuation (3 of 5) of
gb_hcg:AC008749.3 - 119.50 91.09 1.0e+04 110000 ! Continuation (4 of 5) of
gb_hcg:AC026803 - 119.50 85.56 2.1e+04 231450 ! AC026803 Homo sapiens chr
gb_hcg:AP003936 - 117.50 87.45 1.7e+04 138459 ! AP003936 Oryza sativa chr

seq_name: gb_pat:AX031287

seq_documentation_block:
LOCUS AX031287 596 bp DNA linear PAT 20-SEP-2000
DEFINITION Sequence 9 from Patent W09914321.
ACCESSION AX031287
VERSION AX031287.1 GI:10278618

KEYWORDS unclassified.

SOURCE unclassified.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 596)

O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S.,

Huang, D.C. and Strasser, A.

Novel therapeutic molecules

Patent: W0 9914321-A 9 25-MAR-1999;

INST MEDICAL W & E HALL (AU); PUTHALAKATH HANSA (AU); REILLY

LORRAINE O (AU); ADAMS JERRY (AU); CONNOR LIAM O (AU); CORY

SUZANNE (AU); HUANG DAVID C S (AU); STRASSER ANDREAS (AU)

Location/Qualifiers

1..596

/organism="unclassified"

/db_xref="taxon:32644"

1..594

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC09657.1"

/translation="MAKQSDVSVSECDREGLOPAERPPOLRPGAPTSLQTEPQGNP

EGNHPGSGSCPHGSPQGLAPFASPFPATPSLPFIMRSSLRSRSGVFSFDTD

RNPAPKSCDKSTQTPSPQAPFNHYLSANASHMQAEPADMRPEITWIAQELRRIGDEFN

AYAVRRVFLNNTQAAEDHPRMIVILRLRYIVLRVNRMH"

BASE COUNT 145 a 175 c 146 g 130 t

ORIGIN

alignment_scores:

Quality: 1065.00 Length: 198

Ratio: 5.379 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-508-832-10 x AX031287

Align seg 1/1 to: AX031287 from: 1 to: 596

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17

1 ATGGCAAGCAACCTTCGTGATGTAAGTCTGAGTGTGACCGAAGGTAG 50

17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34

51 ACAATGCGAGCTCGGAGAGGCGCTCCCGACCTCAGACCTGGGGCCCTA 100

34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50

101 CCTCCTACAGACAGAGCCACAGGTATCTCTGAAGCAATCAGCGAGGT 150

51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProP 67

151 GAAGGGGACAGCTGCCCCCAGCGAGCGCTCAGGGCCCGCTGGCCCCACC 200

67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84

201 TGGCAGCCCTGGCCCTTTTGCTACACAGATCCCGCTTTTCATCTTTATGA 250

84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100

```

|||||
251 GAAGATCCTCCCTGCTGCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300

101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
301 ACACAGAGGAGCCACACCATGAGTTGTGCACAAATCAACACCAACCCC 350
|||||
117 oSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
|||||
351 AAGTCCCTCTGGCAGGCTTCAACCACTATCTCAGTGCATGCTTCCA 400
|||||
134 eTarGlnAlaGluProAlaAspMetArgProGluIleTPrIleAlaGln 150
|||||
401 TGAGCAGGCTGAACCTGCAGATATGCGCCAGAGATATGATCGCCCA 450
|||||
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgGva 167
|||||
451 GAGTTCGGCGGTATCGGAGACGAGTTTAACTTACTATGCAAGGAGGGT 500
|||||
167 lPheLeuAsnAsnTyrGlnAlaGluAspHisProArgMetValIleL 184
|||||
501 ATTTTGAATAATTACCAAGCAGCCAGACCAACCAACCAATGCTTATCT 550
|||||
184 euArgLeuLeuArgTyrIleValArgLeuValTPrArgMetHis 198
|||||
551 TAGCACTGTACGTTACATTTGCGCGCTGGTGGAGATGCAT 594

seq_name: gb_un:AX031313

seq_documentation_block:
LOCUS AX031313 596 bp DNA linear UNA 20-SEP-2000
DEFINITION Sequence 9 from Patent WO9914321.
ACCESSION AX031313
VERSION AX031313.1 GI:10278641
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,
Huang,D.C. and Strasser,A.
TITLE
Novel therapeutic molecules
PATENT WO 9914321-A 25-MAR-1999;
INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HANSA (AU) ; REILLY
LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY
SUZANNE (AU) ; HUANG DAVID C (AU) ; STRASSER ANDREAS (AU)
FEATURES
Location/Qualifiers
1..596
/organism="unidentified"
/db_xref="taxon:32644"
CDS
1..>594
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09662.1"
/db_xref="GI:10278642"
/translation="MAKQPSDVSSDCREGLOPAERPPOLRPGAPTSLOTPEQGNP
EGNHGSGDSCPHGSPGGLAPASPFPATSPPLFIPMRSSLLSRSSSGYFSDTD
RSPAPGCKSTQTPSPQCAFNHLYLSAMSMROAEADMRPEIWIQAEKLRIGDEFN
AYAVRVELNNYQRAEDHPRMVILRLLYIVLWRKH"
BASE COUNT 145 a 175 c 146 g 130 t
ORIGIN

alignment_scores:
Quality: 1065.00 Length: 198
Ratio: 5.379 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-508-832-10 x AX031313
Align seg 1/1 to: AX031313 from: 1 to: 596

```

```

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
|||||
1 ATGGCAAGCAACCTTCTGATGTAAGTTCTGAGTCTGACCGAGAGGTAG 50
|||||
17 gClnLeuGlnProAlaGluArgProGlnLeuArgProGlyIleAlaProT 34
|||||
51 ACAATTTGACGCTCGGAGAGGCTCTCCAGCTCAGACCTGGGGCCCTA 100
|||||
34 hrSerLeuGlnThrGluProGlnGlnYasnProGluGlyAsnHisGlyGly 50
|||||
101 CCTCCCTACAGACAGAGCCACAAGGTAATCTCAAGGCAATCACGGAGGT 150
|||||
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
|||||
151 GAAGGGGACAGCTGCCCCACGGCAGCCCTCAGGGCCCGCTGGCCCCACC 200
|||||
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
|||||
201 TGCCAGCCCTGGCCCTTTTGTCTACAGATCCCGCTTTTTCATCTTTATGA 250
|||||
84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
|||||
251 GAAGATCCTCCCTGCTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300
|||||
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
301 ACAGACAGGAGCCACGACCCCATGAGTTGTGCACAAATCAACACCAACCCC 350
|||||
117 oSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
|||||
351 AAGTCCCTCTGGCAGGCTTCAACCACTATCTCAGTGCATGCTTCCA 400
|||||
134 eTarGlnAlaGluProAlaAspMetArgProGluIleTPrIleAlaGln 150
|||||
401 TGAGCAGGCTGAACCTGCAGATATGCGCCAGAGATATGATCGCCCA 450
|||||
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgGva 167
|||||
451 GAGTTCGGCGGTATCGGAGACGAGTTTAACTTACTATGCAAGGAGGGT 500
|||||
167 lPheLeuAsnAsnTyrGlnAlaGluAspHisProArgMetValIleL 184
|||||
501 ATTTTGAATAATTACCAAGCAGCCAGACCAACCAACCAATGCTTATCT 550
|||||
184 euArgLeuLeuArgTyrIleValArgLeuValTPrArgMetHis 198
|||||
551 TAGCACTGTACGTTACATTTGCGCGCTGGTGGAGATGCAT 594

```

seq_name: gb_pr:AF032457

seq_documentation_block:

LOCUS AF032457 597 bp mRNA linear PRI 19-FEB-1998
 DEFINITION Homo sapiens BimEL mRNA, complete cds.

ACCESSION AF032457

VERSION AF032457.1 GI:2895495

KEYWORDS

SOURCE human

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 597)

O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,

Cory,S. and Huang,D.C.

Bim: a novel member of the Bcl-2 family that promotes apoptosis

EMBO J. 17 (2), 384-395 (1998)

PUBMED 9430630

2 (bases 1 to 597)

O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,

Cory,S. and Huang,D.C.S.

Direct Submission

Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &

Eliza Hall Institute of Medical Research, PO Royal Melbourne

Hospital, Parkville, Victoria 3050, Australia

FEATURES

source
1..597
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..597

CDS
/note="pro-apoptotic BH3-containing Bcl-2 family member"

146 a 175 c 146 g 130 t
/codon_start=1
/product="BimEL"
/protein_id="AAC39593.1"
/db_xref="GI:2893496"
/translation="MAKOPSDVSECDREGROLOPAERPPOLRCPACPTSLQTEPQGNP
EGNHGEGSDSCPHSGPOGLAPASPGPATSPLEPMRSSLSSSSGVSFSDPD
RSPAPMSCKSTQTPSPPCOENHYLSMASMRQAPADMRPEITHIAQLRRIGDEEN
ATYARVFLNNTYQAEHPRWYLRLLRTIVRLVWRMH"

BASE COUNT 146 a 175 c 146 g 130 t

ORIGIN

alignment_scores:

Quality: 1065.00 Length: 198
Ratio: 5.379 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-508-832-10 x AF032457 ..

Align seg 1/1 to: AF032457 from: 1 to: 597

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
1 ATGGCAAGCAACCTTCTGATGTAAGTCTGAGTCTGACCGAGAGGTAG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProt 34
51 ACRAATTCAGCCTCGGAGAGGCTCCCGAGCTCAGACCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
101 CCTCCCTACAGACAGACCAAGGTAATCTGAAAGGCAATCAGGAGGT 150
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
151 GAAGGGACAGCTGCCCGCCAGCGAGCCCTCAGGCGCCGCTGGCCACC 200
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPhePheMeta 84
201 TGCCAGCCCTGGCCCTTTTGTACCGAGATCCCGCTTTTCATCTTATGA 250
84 rGArgSerSerLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
251 GAAGATCCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTTTTCGAC 300
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
301 ACAGACAGAGCCAGCAGCCCTGAGTGTGACAAATCAACACAAACCCC 350
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
351 AAGTCTCTCTTGGCCAGCGCTTCAACCACTATCTCAGTCGATGGCTTCCA 400
134 eArgGlnAlaGluProAlaAspMetArgProGluIlePrIleAlaGln 150
401 TCAGCGAGCTGAACCTGCAGATATCGCCAGAGATATGATCGCCCAA 450
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgArgVa 167
451 GAGTTCGGCGGTATCGGAGAGGAGTTTACGCTTACTATGCAAGGAGGT 500
167 lPheLeuAsnAsnTyrGlnAlaAlaCluAspHisProArgMetValIleL 184
501 ATTTTGAATAATTACCAAGCAGCGGAGACCAACCCAGCAATGGTTATCT 550

184 euArgLeuLeuArgTyrIleValargLeuValTrpArgMetHis 198
551 TAGCACTGTTACGTTACATTCCTGCCCTGGTGGAGAAATGCAT 594

seq_name: gb_pr:AB071197

seq_documentation_block:

LOCUS AB071197 630 bp mRNA linear PRI 05-DEC-2001
DEFINITION Homo sapiens mRNA for Bim-betal, complete cds.
ACCESSION AB071197
VERSION AB071197.1 GI:17351900
KEYWORDS
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens

REFERENCE
1 (sites)
U.M., Miyashita,T., Shikama,Y., Tadokoro,K. and Yamada,M.
Molecular cloning and characterization of six novel isoforms of
human Bim, a member of the proapoptotic Bcl-2 family(1)
FEBS Lett. 509 (1), 135-141 (2001)

PUBMED 11734221

REFERENCE 2 (bases 1 to 630)

AUTHORS Miyashita,T.

TITLE Direct Submission

JOURNAL Submitted (06-SEP-2001) Toshiyuki Miyashita, National Children's
Medical Research Center, Department of Genetics, 3-35-31 Taishido,
Setagaya, Tokyo 154-8509, Japan (E-mail:tmlyashita@nch.go.jp,
Tel:81-3-3414-8121(ex.2763), Fax:81-3-3414-3208)

FEATURES
Location/Qualifiers

1..630
/organism="Homo sapiens"

/db_xref="taxon:9606"

1..408
/gene="bim-betal"

1..408
/gene="bim-betal"

/product="Bim-betal"

/protein_id="BAB78591.1"

/db_xref="GI:17351901"

/translation="MAKOPSDVSECDREGROLOPAERPPOLRCPACPTSLQTEPQGNP
EGNHGEGSDSCPHSGPOGLAPASPGPATSPLEPMRSSLSSSSGVSFSDPD
RSPAPMSCKSTQTPSPPCOENHYLSMASMRQAPADMRPEITHIAQLRRIGDEEN"

BASE COUNT 157 a 186 c 152 g 135 t

ORIGIN

alignment_scores:

Quality: 1049.50 Length: 209
Ratio: 5.301 Gaps: 1
Percent Similarity: 94.737 Percent Identity: 94.737

alignment_block:

US-09-508-832-10 x AB071197 ..

Align seg 1/1 to: AB071197 from: 1 to: 630

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
1 ATGGCAAGCAACCTTCTGATGTAAGTCTGAGTCTGACCGAGAGGTAG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProt 34
51 ACRAATTCAGCCTCGGAGAGGCTCCCGAGCTCAGACCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
101 CCTCCCTACAGACAGACCAAGGTAATCTGAAAGGCAATCAGGAGGT 150
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67

```

|||||
151 CAAGGGACAGCTGCCCCACGGACGCTCAGGGCCCGCTGGCCACC 200
|||||
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
|||||
201 TCCAGGCCCTGGCCCTTTGCTACCAAGATCCCGCTTTCATCTTTATGA 250
|||||
84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrrPheSerPheAsp 100
|||||
251 GAAGATCCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTCTTTGAC 300
|||||
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
301 ACACAGAGGAGCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350
|||||
117 oSerProProCysGlnAlaPheAsnHisTyrrLeuSerAlaMet..... 131
|||||
351 AAGTCCCTCTGCCAGGCCCTTCAACCACTATCTCAGTCCCAATGGCTAACT 400
|||||
132 .....AlaSerMetArgGlnAlaGluPro 139
|||||
401 GGGACTAGAAACAGCTCCATCCAGGCTTCATGAGCGAGGCTGAACCT 450
|||||
140 AlaAspMetArgProGluIleThrPheAlaGlnGluLeuArgGlyIleG 156
|||||
451 GCACATATGCCCCCAGAGATATGATGATGATGATGATGATGATGATGAT 500
|||||
156 yAspGluPheAsnAlaTyrrTyrrAlaArgArgValPheLeuAsnAsnTyrg 173
|||||
501 AGACGAGTTTAAAGCTTACTATGCAAGGAGGGGTATTTTTCGAATATATAC 550
|||||
173 lAlaAlaGluAspHisProArgMetValIleLeuArgGluLeuArgTyrr 189
|||||
551 AAGCAGCGGAAGACACCCACCAATGTTATCTTACGACTGTTAGGTAC 600
|||||
190 lIleValArgLeuValTrpArgMetHis 198
|||||
601 ATTGTCCGCTGGTGTGGAGATGATGAT 627

```

seq_name: 9b_pr:AB071195

seq_documentation_block:

LOCUS AB071195 722 bp mRNA linear PRI 05-DEC-2001
 DEFINITION Homo sapiens mRNA for Bim-alpha1, complete cds.
 ACCESSION AB071195
 VERSION AB071195.1 GI:17351896
 KEYWORDS
 SOURCE Homo sapiens CDNA to mRNA.
 ORGANISM

REFERENCE
 AUTHORS
 TITLE U.M., Miyashita.T., Shikama.Y., Tadokoro.K. and Yamada.M.
 Molecular cloning and characterization of six novel isoforms of
 human Bim, a member of the proapoptotic Bcl-2 family(1)
 JOURNAL FEBS Lett. 509 (1), 135-141 (2001)
 PUBMED 11734221
 REFERENCE 2 (bases 1 to 722)
 AUTHORS Miyashita.T.
 TITLE Direct Submission

JOURNAL Submitted (06-SEP-2001) Toshiyuki Miyashita, National Children's
 Medical Research Center, Department of Genetics; 3-35-31 Taishido,
 Setagaya, Tokyo 154-8509, Japan (E-mail:umiya@taishido.nich.go.jp,
 Tel:81-3-3414-8121(ex.2763), Fax:81-3-3414-3208)
 FEATURES
 source
 1..722
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene
 1..510
 /gene="bim-alpha1"
 CDS
 1..510
 /gene="bim-alpha1"
 /note="alternatively spliced: see also Acc#

```

AB071196-AB071200*
/codon_start=1
/product="bim-alpha1"
/protein_id="BAB78589.1"
/db_xref="GI:17351897"
/translation="MAKOPSDVSECDREGLOAERPPOLRGPATSLQTEPOGNP
EGNHGGEGDSCPHGSPGPIATSPFATRSPPLFIMRRSSLLSRSSGFFSFDTD
RSPAPHCDDKSTQTPSPPCQAFNHILSAMASHMQAEPADMRPEIWIAGELRRIGDEFN
AYIARRLEK"
BASE COUNT 180 a 190 c 193 g 159 t
ORIGIN

```

alignment_scores:
 Quality: 1010.50 Length: 240
 Ratio: 5.104 Gaps: 2
 Percent Similarity: 82.500 Percent Identity: 82.083

alignment_block:

US-09-508-832-10 x AB071195 ..

Align seg 1/1 to: AB071195 from: 1 to: 722

```

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
|||||
1 ATGCCAAAGCAACCTTCTGATGAAGTTCTGAGTGTGACCGAGAGGTAG 50
|||||
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
|||||
51 ACAATTCGAGCTCGCGAGAGGCTCCCGAGCTCAGACCTGGGGCCCTA 100
|||||
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
|||||
101 CCTCCCTACACAGAGCCACAGGTAATCTCGAGGCAATCACGGAGGT 150
|||||
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
|||||
151 GAAGGGAGACGCTGCCCGACGGCCTCAGGGCCCGCTGGGGCCACC 200
|||||
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
|||||
201 TGCCAGACCTGGCCCTTTTGTCTACCAAGATCCCGCTTTTCATCTTTATGA 250
|||||
84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrrPheSerPheAsp 100
|||||
251 GAAGATCCTCCCTGCTGCTCGATCTCCAGTGGGTATTTCTCTTTGAC 300
|||||
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
|||||
301 ACAGACAGGAGCCAGCACCCATGAGTTGTGACAAATCAACACAAACCCC 350
|||||
117 oSerProProCysGlnAlaPheAsnHisTyrrLeuSerAlaMetAlaSerM 134
|||||
351 AAGTCCCTCTTCCAGGCCCTTCAACCACTATCTCAGTCCCAATGGCTTCCA 400
|||||
134 eArgGlnAlaGluProAlaAspMetArgProGluIleThrPheAlaGln 150
|||||
401 TGAGCGAGGCTGAACCTGCGAGATATGCGCCAGAGATATGGATCGCCCA 450
|||||
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrrTyrrAla..... 164
|||||
451 GAGTTGCGGGCGTATCGGAGACGAGTTTAAACGCTTACTATGCAAGAGGTT 500
|||||
164 .....Arg..ArgValPheLeuAsnAsnTyrrGlnAlaAla 175
|||||
501 ACAGAAATAGAGGAAGTTGTCGTGTAGTTCTCATGTTATTCAGTCCACCTTA 550
|||||
164 ..... 164
551 AGCGCAGCTGGGGAAGCGCTTTTCAGACCGGAGCGCTGCGAGCGCTCAATCCTTGA 600
|||||
165 .....Arg..ArgValPheLeuAsnAsnTyrrGlnAlaAla 175
|||||

```

seq_name: gb_pat:AX031283	seq_documentation_block:	LOCUS	AX031283	590 bp	DNA	linear	PAT 20-SEP-2000
DEFINITION	Sequence 5 from Patent WO9914321.	LOCUS	AX031283	590 bp	DNA	linear	PAT 20-SEP-2000
ACCESSION	AX031283	VERSION	AX031283.1	GI:10278614			
KEYWORDS	unidentified.	ORGANISM	unidentified.				
SOURCE	unclassified.	REFERENCE	1 (bases 1 to 590)				
AUTHORS	O'Reilly,L., Puthalakath,H., Adams,J., O'Connor,L., Cory,S.,	AUTHORS	Huang,D.C. and Strasser,A.				
TITLE	Novel therapeutic molecules	TITLE	INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HAMSA (AU) ; REILLY				
JOURNAL	LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY	JOURNAL	SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)				
FEATURES	Location/Qualifiers	FEATURES	1..590				
source	/organism="unidentified"	source	/db_xref="taxon:32644"				
CDS	1..>587	CDS	/notes="unnamed protein product"				
	/codon_start=1		/protein_id="CAC09655.1"				
	/db_xref="GI:10278615"		/translation="NAKOPSDVSSCEDREGGLOPAERPPOLRPGAPTSIQTPEQNP				
	DGEGDRCPHGSPOGGLAPSPGFPATRSPLFIIVRRSSLLSSSSGYSFDFDRSPA		PMSCDKSTOTPPPCQAENHYLSAMASIRQSEPEDELRLPEIRIAQLRIGDEFNET				
	YTRRVFANDYREAEHPQVLIQLRFRFLVRR"						
BASE COUNT	137 a 178 c 150 g 125 t	BASE COUNT	137 a 178 c 150 g 125 t				
ORIGIN		ORIGIN					
alignment_scores:	Quality: 908.00	alignment_scores:	Quality: 908.00				
	Ratio: 4.935		Ratio: 4.935				
	Percent Similarity: 92.000		Percent Similarity: 92.000				
alignment_block:	US-09-508-832-10 x AX031283	alignment_block:	US-09-508-832-10 x AX031283				
Align seg 1/1 to: AX031283 from: 1 to: 590		Align seg 1/1 to: AX031283 from: 1 to: 590					
1 MetAlaLysGlnProSerAspValSerSerGluCysaspargGluGlyar 17		1 MetAlaLysGlnProSerAspValSerSerGluCysaspargGluGlyar 17					
1 ATGGCCACCAACCTTCGTAGTGAAGTTCCTCAGTGTGACAGAGAAGGTGG 50		1 ATGGCCACCAACCTTCGTAGTGAAGTTCCTCAGTGTGACAGAGAAGGTGG 50					
17 gGlnLeuGlnProAlaGluArgProToGlnLeuArgProGlyAlaPro 34		17 gGlnLeuGlnProAlaGluArgProToGlnLeuArgProGlyAlaPro 34					
51 ACAATTGCAGCTGTGTGAGAGGCGCTCCCGACCTCAGGCGCTGGGCCCCCTA 100		51 ACAATTGCAGCTGTGTGAGAGGCGCTCCCGACCTCAGGCGCTGGGCCCCCTA 100					
34 hrSerLeuGlnThrGluProGlnGlyAenProGluGlyAsnHlsGlyGly 50		34 hrSerLeuGlnThrGluProGlnGlyAenProGluGlyAsnHlsGlyGly 50					
101 CCTCCCTACAGACAGACGACCGCAGGTAATCCGAC.....GGC 138		101 CCTCCCTACAGACAGACGACCGCAGGTAATCCGAC.....GGC 138					
51 GluGlyAspSerCysProHlsGlySerProGlnGlyProLeuAlaProPr 67		51 GluGlyAspSerCysProHlsGlySerProGlnGlyProLeuAlaProPr 67					
139 GAAGGGGACCGGTGCCCCCAGCGAGCGCTCAGGCGCGCTGGCCCCACC 198		139 GAAGGGGACCGGTGCCCCCAGCGAGCGCTCAGGCGCGCTGGCCCCACC 198					
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheillePheMeta 84		67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheillePheMeta 84					

US-09-508-832-10 x AX031309 ..

Align seg 1/1 to: AX031309 from: 1 to: 590

```

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
1 ATGGCCAGGACCACTTCTGATGTAAGTCTCAGTCTGACAGAGAGGTGG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
51 ACAATTGACGCTGCTGAGAGGCTCCCGAGCTCAGGCTGGGCCCTTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGly 50
101 CCTCCCTACAGACAGACCAAGTAATCCGAC.....GGC 138
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPro 67
139 GAAGGGGACCGCTGCCCGCCACGCGAGCTCAGGCGCCGCTGGGCCAC 188
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMe 84
189 GCCCAGGCTGGCCCTTTGCTACAGATCCCACTTTTCATCTTTGTGA 238
84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
239 GAAGATCTTCTGCTGTCGCTCCGCTCCTCAGTGGGTATTTCTCTTT 288
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThr 117
289 ACAGACAGGAGCGCGCACCATGAGTTGTGACAAGTCAACACAAACCC 338
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSe 134
339 AAGTCTCTTGGCAGGCTTCAACCACTATCTCAGTGCATGGCTTCCA 388
134 etArgGlnAla.....GluProAlaAspMetArgProGluIleTrp 148
389 TACGACAGTCTCAGGAGGACCTGAGATCTGCCCGCGGATACGGATT 438
149 AlaGlnLeuLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAl 165
439 GCACAGGAGCTGGCGGGATCGGAGACGAGTTCACAGAACTTACACA 488
165 gArgValPheLeuAsnAspTyrGlnAlaAlaGluAspHisProArgMe 182
489 GAGGCTCTTTGCAATGATTACCGGAGGCTGAAGACCACTCAATGG 538
182 alIleLeuArgLeuLeuArgTyrIleValArgLeuValTyrArgMeth 198
539 TTATCTTACAACTGTACGCTTATCTTCCGCTGCTGATGAGAAGCAT 588

```

seq_name: gb_ro:AF032459

seq_documentation_block:

LOCUS AF032459 591 bp mRNA linear ROD 19-FEB-1998

DEFINITION Mus musculus BimEL mRNA, complete cds.

ACCESSION AF032459

VERSION AF032459.1 GI:2895499

KEYWORDS house mouse.

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 591)

O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M.,

Cory, S. and Huang, D.C.

Bim, a novel member of the Bcl-2 family that promotes apoptosis

EMBO J 17 (2), 384-395 (1998)

98094360

PUBLISHED 9430630

REFERENCE 2 (bases 1 to 591)

O'Connor, L., Strasser, A., O'Reilly, L.A., Hausmann, G., Adams, J.M.,

Cory, S. and Huang, D.C.S.

Direct Submission

Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter &

Eliza Hall Institute of Medical Research, PO Royal Melbourne

Hospital, Parkville, Victoria 3050, Australia

Location/Qualifiers

1..591

/organism="Mus musculus"

/db_xref="taxon:10090"

/note="pro-apoptotic Bcl-2 family member"

/codon_start=1

/product="BimEL"

/protein_id="AAC40029.1"

/db_xref="GI:2895500"

/translation="MAKOPSDVSSCEDREGGQLQAPRRPQLRPGAPTSLQTEPQGNP

DGEGDRCPHGSPQGLAPFPAPSPGFATRSPLFIFVRRSLLSRSSSGYFSDTRSPA

PKSCDKSTQTFSPPCQAFNFHLSAMASIRQSEPEDLRPEIRIAQLRRIQDGFNET

YTRRVFANDYREADHPQWILQLRLEIFRLVRRH"

BASE COUNT 138 a 178 c 150 g 125 t

ORIGIN

alignment_scores:

Quality: 908.00 Length: 200

Ratio: 4.935 Gaps: 2

Percent Similarity: 92.000 Percent Identity: 86.500

alignment_block:

US-09-508-832-10 x AF032459 ..

Align seg 1/1 to: AF032459 from: 1 to: 591

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17

1 ATGGCCAGGACCACTTCTGATGTAAGTCTCAGTCTGACAGAGAGGTGG 50

17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34

51 ACAATTGACGCTGCTGAGAGGCTCCCGAGCTCAGGCTGGGCCCTTA 100

34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGly 50

101 CCTCCCTACAGACAGACCAAGTAATCCGAC.....GGC 138

51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPro 67

139 GAAGGGGACCGCTGCCCGCCACGCGAGCTCAGGCGCCGCTGGGCCAC 188

67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMe 84

189 GCCCAGGCTGGCCCTTTGCTACAGATCCCACTTTTCATCTTTGTGA 238

84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100

239 GAAGATCTTCTGCTGTCGCTCCGCTCCTCAGTGGGTATTTCTCTTT 288

101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThr 117

289 ACAGACAGGAGCGCGCACCATGAGTTGTGACAAGTCAACACAAACCC 338

117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSe 134

339 AAGTCTCTTGGCAGGCTTCAACCACTATCTCAGTGCATGGCTTCCA 388

134 etArgGlnAla.....GluProAlaAspMetArgProGluIleTrp 148

389 TACGACAGTCTCAGGAGGACCTGAGATCTGCCCGCGGATACGGATT 438

149 AlaGlnLeuLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAl 165

439 GCACAGGAGCTGGCGGGATCGGAGACGAGTTCACAGAACTTACACA 488

165 gArgValPheLeuAsnTyrGlnAlaGluAspHisProArgMetV 182
 489 GAGGCTGTTGCAATGATTACCGGAGCTGAAGACACCCCTCAATGG 538
 182 alileuArgLeuLeuArgTyrIleValArgLeuValTyrArgMetHis 198
 539 TTATCTTAACTGTTACGCTTATCTCCGCTGTGATGGAAGAGCAT 588

seq_name: gb_ro:BC019556

seq_documentation_block:
 LOCUS BC019556 1137 bp mRNA linear ROD 22-JAN-2002
 DEFINITION Mus musculus, BCL2-like 11 (apoptosis facilitator), clone MGC:28730
 IMAGE:4459720, mRNA, complete cds.
 ACCESSION BC019556
 VERSION BC019556.1 GI:18043532
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1137)
 DIRECT SUBMISSION
 SUBMITTED (19-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC project URL: http://mgc.nci.nih.gov
 CONTACT: MGC help desk
 EMAIL: cgabps-remail.nih.gov
 TISSUE: Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA LIBRARY PREPARATION: Life Technologies, Inc.
 DNA LIBRARY ARRANGED BY: The I.M.A.G.E. Consortium (LIML)
 DNA SEQUENCING BY: Baylor College of Medicine Human Genome
 Sequencing Center
 CENTER CODE: BCM-HGSC
 WEB SITE: http://www.hgsc.bcm.tmc.edu/cdna/
 CONTACT: amg@bcm.tmc.edu
 GUNARATNE, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 YOON, V.S., KOWIS, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT NIH-MGC project URL: http://mgc.nci.nih.gov
 CONTACT: MGC help desk
 EMAIL: cgabps-remail.nih.gov
 TISSUE: Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA LIBRARY PREPARATION: Life Technologies, Inc.
 DNA LIBRARY ARRANGED BY: The I.M.A.G.E. Consortium (LIML)
 DNA SEQUENCING BY: Baylor College of Medicine Human Genome
 Sequencing Center
 CENTER CODE: BCM-HGSC
 WEB SITE: http://www.hgsc.bcm.tmc.edu/cdna/
 CONTACT: amg@bcm.tmc.edu
 GUNARATNE, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 YOON, V.S., KOWIS, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov
 Series: IRAK plate: 37 Row: p Column: 14.

FEATURES

source
 1. .1137
 /organism="Mus musculus"
 /db_xref="LocusID:12125"
 /db_xref="taxon:10090"
 /clone="MGC:28730 IMAGE:4459720"
 /tissue_type="Mammary tumor. Metallothionin-TGF alpha
 model. 10 month old virgin mouse. Taken by biopsy."
 /clone_lib="NCL CGAP_Mam1"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 223. .813
 /codon_start=1
 /product="BCL2-like 11 (apoptosis facilitator)"
 /protein_id="AAH19556.1"
 /db_xref="GI:18043533"
 /translation="MAKQPSDVSECDREGGQLQAPRPPQLRPGAPTSLQTEPQGNP
 DGGDRCPHGSPPQGLFAPRSPFATRSPLFFVRRSLLSRSSGYSFSDTDRSPA
 PMSCKSTQTPSPCOAFNPVLSAMSIROSEPEDLRPEIRIAQLRRIGDEFNET
 VYRRVFANDYREADHPQWVILQLLEIFLVVRRH"

CDS

BASE COUNT 275 a 330 c 282 g 250 t
 ORIGIN
 alignment_scores:
 Quality: 908.00 Length: 200
 Ratio: 4.935 Gaps: 2

Percent Similarity: 92.000 Percent Identity: 86.500
 alignment_block:
 US-09-508-832-10 x BC019556
 Align seg 1/1 to: BC019556 from: 1 to: 1137
 1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
 223 ATGCCCAAGCAACCTTCATGATGAAGTCTGAGTGTGACAGAGAAGTGG 272
 17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
 273 ACAATTCAGACCTGCTGAGAGGCTCTCCAGCTCAGGCTGGGGGCCCTA 322
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
 323 CCTCCTACAGACAGAACCGCAAGGTAAATCCCGAC.....GGC 360
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
 361 GAAGGGGACCGCTGCCCCACCGGAGCCCTCAGGGCCCGCTGGCCGCCACC 410
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
 411 GCCACGCCCTGGCCCTTTGCTACCATGCCACATTTTCATCTTTGTGA 460
 84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
 461 GAAGATCTTCTCTGCTGCCCGCTCCAGTGGGTATTTCTCTTTTGAC 510
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
 511 ACAGACAGAGAGCCCGGACCCATGAGTTGTGACAGTCAACACAAACCC 560
 117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
 561 AAGTCTCTCTGCCAGGCTTCAACCACTATCTCAGTGCATGGCTTCCA 610
 134 etArgGlnAla.....GluProAlaAspMetArgProGluIleTrrpIle 148
 611 TAGCAGAGCTCAGGAGGAACCTCAAGATCTCGCCCGGAGATACGGATT 660
 149 AlaGlnGluLeuArgArgTleGlyAspGluPheAsnAlaTyrTyrAlaAr 165
 661 GCACAGAGCTCGCGCGATCGGAGACGAGTTCACGCAAACTTACACAG 710
 165 gArgValPheLeuAsnAsnTyrGlnAlaGluAspHisProArgMetV 182
 711 GAGGGTGTGTTGCAATGATTACCGGAGGCTGAAGAGCACCCCTCAAATGG 760
 182 alileuArgLeuLeuArgTyrIleValArgLeuValTyrArgMetHis 198
 761 TTATCTTAACTGTTACGCTTATCTTCGCTGTGATGGAAGAGCAT 810

seq_name: gb_ro:AF065433

seq_documentation_block:

LOCUS AF065433 591 bp mRNA linear ROD 11-MAR-1999
 DEFINITION Rattus norvegicus Bcl-2 related ovarian death gene product BOD-2
 mRNA, complete cds.
 ACCESSION AF065433
 VERSION AF065433.1 GI:3228569
 KEYWORDS Norway rat.
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 591)
 Hsu, S.Y., Lin, P., and Hsueh, A.J.
 BOD (Bcl-2-related ovarian death gene) is an ovarian BH3
 domain-containing proapoptotic Bcl-2 protein capable of

```

dimerization with diverse antiapoptotic Bcl-2 members
98400436
REFERENCE
MEDLINE
AUTHORS
TITLE
JOURNAL
SUBMITTED (15-MAY-1998) GYN/OB, Stanford University, MSOB S385,
Stanford, CA 94305, USA
FEATURES
Location/Qualifiers
1..591
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1..591
/codon_start=1
/product="Bcl-2 related ovarian death gene product BOD-L"
/protein_id="AAC23595.1"
/db_xref="GI:3228570"
/translation="MAKQSDVNSDDEGQLOPAERPPOLRPGAPTSLOTESQGNP
DEGDRCPHSGPGLAPASPGFATRSPLFIEVRRSSLSRSSSGYFSFDTRSPA
PHSCDKSTPTSPPCQAFNHYSAMASTROSEPEEDLPEIRIAQLRRIGDEFNET
YTRAFANDYREADHPQMVILQLLRFRLVRRH"
BASE COUNT 139 a 177 c 153 g 122 t
ORIGIN

alignment_scores:
Quality: 893.00 Length: 200
Ratio: 4.880 Gaps: 2
Percent Similarity: 91.500 Percent Identity: 85.000

alignment_block:
US-09-508-832-10 x AF065433
Align seg 1/1 to: AF065433 from: 1 to: 591

1 MetAlaLysGlnProSerAspValSerGluCysAspArgGluGlyAr 17
1 ATGCCAAGCAACCTCTGATGTAATCTGAGTGTACAGAGAGGTGG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPro 34
51 ACAATTGACCTGCTGACAGAGCCCTCCCACTGAGCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
101 CCTCCCTACAGACAGAAATCGCAAGGTAAATCCCGAC.....GGC 138
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
139 GAAGGGACCCGCTGCCCAACGACAGCCCTCAGGGCCCGCTGGCCACC 188
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
189 GCCCAGCCCTGGTCTTTGGTACCAGATCCCCACTTTTCATCTTTGTGA 238
84 rGArgSerSerLeuLeuSerArgSerSerGlyTyPheSerPheAsp 100
239 GAAGATCTTCTGCTGCTGCCGCTCTCCAGTGGGTATTTCTCTTTGAC 288
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
289 ACAGACAGGAGCCCGCCACCATGATGCTGACAACTCACACACACCC 338
117 oSerProProCysGlnAlaPheAsnHisTyPheSerAlaMetaSerM 134
339 AAGTCCCTCTTGGCAGCCCTTCAACCATATATCTCAGTGCATGCTTCCA 388
134 eTArgGlnAla.....GluProAlaAspMetArgProGluIleTyrPile 148
389 TAAGCCAGTCTCAGGAGGAACCTGAAGATCTGCCGCCAGACATACGCATC 438
149 AlaGlnLeuArgArgIleGlyAspGluPheAsnAlaTyrTyAlaAr 165
|||||

```

```

439 GCACAGGAGCTGCGCGGAGCGAGAGCTCAATGAGACTTACACGAG 488
165 eArgValPheLeuAsnAspTyrGlnAlaAlaGluAspHisProArgMetV 182
|||||
489 GAGGCGCTTTCGAACGATTACCGAGGCGGAGACACCCCAATGG 538
182 aLlLeuArgLeuLeuArgTyrIleValArgLeuValTrpArgMetHis 198
|||||
539 TTATCTTAAACTGTTACGATTCATCTTCGTCGTCGTGAGAGAGGAC 588
seq_name: gb_pr:AB071198

seq_documentation_block:
LOCUS AB071198 493 bp mRNA linear PRI 05-DEC-2001
DEFINITION Homo sapiens mRNA for Bim-beta2, complete cds.
ACCESSION AB071198
VERSION AB071198.1 GI:17351902
KEYWORDS
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
AUTHORS U.M. Miyashita,T., Shikama,Y., Tadokoro,K. and Yamada,M.
TITLE Molecular cloning and characterization of six novel isoforms of
human Bim, a member of the Proapoptotic Bcl-2 family(1)
JOURNAL FEBS Lett. 509 (1), 135-141 (2001)
PUBMED 11734221
REFERENCE 2 (bases 1 to 493)
AUTHORS Miyashita,T.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2001) Toshiyuki Miyashita, National Children's
Medical Research Center, Department of Genetics; 3-35-31 Taisshido,
Setagaya, Tokyo 154-8509, Japan (E-mail:tmivashita@nch.go.jp,
Tel:81-3-3414-8121(ex-2763), Fax:81-3-3414-3208)
FEATURES
source
1..493
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..408
/gene="bim-beta2"
1..408
/gene="bim-beta2"
/feature="bim-beta2"
/translation="MAKQSDVNSDDEGQLOPAERPPOLRPGAPTSLOTESQGNP
EGNHGEGDSCPHSGPGLAPASPGFATRSPLFIEVRRSSLSRSSSGYFSFDTRSPA
PHSCDKSTPTSPPCQAFNHYSAMASTROSEPEEDLPEIRIAQLRRIGDEFNET
YTRAFANDYREADHPQMVILQLLRFRLVRRH"
BASE COUNT 119 a 151 c 114 g 109 t
ORIGIN

alignment_scores:
Quality: 837.50 Length: 199
Ratio: 5.138 Gaps: 2
Percent Similarity: 81.910 Percent Identity: 81.407

alignment_block:
US-09-508-832-10 x AB071198
Align seg 1/1 to: AB071198 from: 1 to: 493

1 MetAlaLysGlnProSerAspValSerGluCysAspArgGluGlyAr 17
1 ATGCCAAGCAACCTCTGATGTAATCTGAGTGTACAGAGAGGTGG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPro 34
51 ACAATTGACCTGCTGACAGAGCCCTCCCACTGAGCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
101 CCTCCCTACAGACAGAAATCGCAAGGTAAATCCCGAC.....GGC 138
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
139 GAAGGGACCCGCTGCCCAACGACAGCCCTCAGGGCCCGCTGGCCACC 188
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
189 GCCCAGCCCTGGTCTTTGGTACCAGATCCCCACTTTTCATCTTTGTGA 238
84 rGArgSerSerLeuLeuSerArgSerSerGlyTyPheSerPheAsp 100
239 GAAGATCTTCTGCTGCTGCCGCTCTCCAGTGGGTATTTCTCTTTGAC 288
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
289 ACAGACAGGAGCCCGCCACCATGATGCTGACAACTCACACACACCC 338
117 oSerProProCysGlnAlaPheAsnHisTyPheSerAlaMetaSerM 134
339 AAGTCCCTCTTGGCAGCCCTTCAACCATATATCTCAGTGCATGCTTCCA 388
134 eTArgGlnAla.....GluProAlaAspMetArgProGluIleTyrPile 148
389 TAAGCCAGTCTCAGGAGGAACCTGAAGATCTGCCGCCAGACATACGCATC 438
149 AlaGlnLeuArgArgIleGlyAspGluPheAsnAlaTyrTyAlaAr 165
|||||

```


34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
 101 CTTCTCTACACAGACAGCCACAGGTAACTCTGAAGCCAAATCAGGAGT 150
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
 151 GAAGGGGACAGCTGCCCCACGACGCCCTCAGGGCCCGCTGGCCCCACC 200
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMet 84
 201 TGCCAGCCCTGGCCCTTTTGTCTACAGATCCCGCTTTTCATCTTTATGA 250
 84 rArgSerSerLeuLeuArgSerSerSerClyTyrrPheSerPheAsp 100
 251 GAAGATCTCTCTGCTGCTGATCTCCAGTGGGTATTTCTTTTTCGAC 300
 101 ThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnThrPr 117
 301 ACAGACAGAGCCAGCCAGCCATGAGTGTGACAAATCAACAAACCC 350
 117 oSerProProCysGlnAlaPheAsnHisTyrrLeuSerAlaMetAlaSer 133
 351 AGTCTCTCTGCCAGCCCTCAACCACTATCTCAGTGCATGGGTA... 397
 134 MetArgGlnAlaGluProAlaAspMetArgProGluIleTrpIleAlaG 150
 397 397
 150 nGluLeuArgArgIleGlyAspGluPheAsnAlaTyrrAlaArgv 167
 397 397
 167 alphaPheAsnAspTyrrClnAlaAlaClnAspHisProArgMetValIle 183
 398 ..TTTITGAAATTTCCACAGCCAGCCAGACCCACCAAGTGTATTC 445
 184 LeuArgLeuLeuArgTyrrIleValArgLeuValTrpArgMetHis 198
 446 TTACGACTGTACGTATGATTCCTGCTGGTGGAGAAATGCAT 490

seq_name: gb_pr:AC096670

seq_documentation_block: 181497 bp DNA linear PRI 19-SEP-2001
 LOCUS AC096670
 DEFINITION Homo sapiens chromosome 2 clone RP11-438K19, complete sequence.
 ACCESSION AC096670 AC013332
 VERSION AC096670.1 GI:15668150
 KEYWORDS HTG.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 181497)

Waterston, R.H.

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 181497)

Waterston, R.H.

Direct Submission

Submitted (19-SEP-2001) Genome Sequencing Center, Washington

University School of Medicine, 444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Sep 19, 2001 this sequence version replaced gi:7248987.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@genome.wustl.edu

----- Project Information -----

Center project name: H_NH0438K19

Drafting center: WIBR

FEATURES
 source

Location/Qualifiers

1..181497

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="2"

/clone="RP11-438K19"

BASE COUNT 52329 a 40542 c 39097 g 49529 t

ORIGIN

alignment_scores:

Quality: 713.00 Length: 131

Ratio: 5.443 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-508-832-10 x AC096670/rev ..

Align seg 1/1 to reverse of: AC096670 from: 1 to: 181497

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17

75475 ATGGCAAGCAACCTTCTGATGTAAGTTCAGTGTGACCGAGAAGGTAG 75426

17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34

75425 ACAATTGCGAGCTGCGGAGAGGCTTCCCGACGCTCAGACCTGCGGCCCTTA 75376

34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50

75375 CCTCCCTACAGACAGAGCCACAAAGCTAATCTCTGAAGCAATCAGGAGCT 75326

51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67

75325 GAAGGGGACAGCTGCCCCACGCGAGCCCTCAGGGCCCGCTGGCCCCACC 75276

67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMet 84

75275 TGCAGCCCTGGCCCTTTTGTCTACCAAGATCCCGCTTTTCATCTTTATGA 75226

84 rArgSerSerLeuLeuSerArgSerSerSerSerGlyTyrrPheSerPheAsp 100

75225 GAAGATCTCTCTGCTGCTCGATCTCTCAGTGGGTATTTCTCTTTGAC 75176

101 ThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnThrPr 117

75175 ACAGACAGAGCCAGCCAGCCATGAGTGTGACAAATCAACAAACCC 75126

117 oSerProProCysGlnAlaPheAsnHisTyrrLeuSerAlaMet 131

75125 AAGTCTCTCTGCGAGGCTTCAACCACTATCTCAGTGCATG 75083

seq_name: gb_pat:AX031285

seq_documentation_block:

LOCUS AX031285

DEFINITION Sequence 7 from Patent WO9914321.

ACCESSION AX031285

VERSION AX031285.1 GI:10278616

KEYWORDS

SOURCE

unidentified.

ORGANISM

unclassified.

REFERENCE

1 (bases 1 to 416)

O'Reilly, L., Puthalakath, H., Adams, J., O'Connor, L., Cory, S.,

Huang, D.C. and Strasser, A.

Novel therapeutic molecules

Patent: WO 9914321-A 7 25-MAR-1999;

INST MEDICAL W & E HALL (AU) ; PUTHALAKATH HANSA (AU) ; REILLY

LORRAINE O (AU) ; ADAMS JERRY (AU) ; CONNOR LIAM O (AU) ; CORY

SUZANNE (AU) ; HUANG DAVID C S (AU) ; STRASSER ANDREAS (AU)

Location/Qualifiers

FEATURES

```

source          1. 416
                /organism="unidentified"
                /db_xref="taxon:32644"
                1. >414
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAC09656.1"
                /db_xref="GI:10278617"
                /translation="MAKQPSDVSSECDREGROLOPAERPPOLRPGAPTSIQTEPODORS
                PAPMSCDKSTQTPSPPCQAFNHVLSAMASMRQAEADMRPEIWIQAQLRIGDEFNAY
                YARRVFLNNTQAEEDHPRMVILRLRYIVRLVRMH"
BASE COUNT      113 a 113 c 103 g 87 t
ORIGIN
alignment_scores:
  Quality: 696.00      Length: 198
  Ratio: 5.043        Gaps: 1
  Percent Similarity: 69.697   Percent Identity: 69.697
alignment_block:
US-09-508-832-10 x AX031285 ..
Align seg 1/1 to: AX031285 from: 1 to: 416
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
1 ATGGCAAGGCAACCTCTGATGTAAGTCTGAGTGTGACCGAGAAGGTAG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProt 34
51 ACAATTGCAGCTCGCGAGAGGCGCTCCACGCTCAGACCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
101 CCTCCTACAGACAGAGCCACAA..... 123
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
123 ..... 123
67 oAlaSerProGlyPropheAlaThrArgSerProLeuPheIlePheMeta 84
123 ..... 123
84 rArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
123 ..... 123
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
124 ...GACAGGAGCCGACGCCATGAGTTGTGACAAATCAACACAAACCCC 170
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
171 AAGTCCTCCTTGCAGCGCTTCAACCACTATCTCAGTGCATGCTTCCA 220
134 eArgGlnAlaGluProAlaAspMetArgProGluIleTrrPileAlaGln 150
221 TGAGCAGGCTGAACCTTCAGATATGCGCCAGAGATATGATGCGCCAA 270
151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgArgVa 167
271 CAGTTGGCGGTATCGGAGACGAGTTTACGCTTACTATGCAAGGAGGT 320
167 lPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetValIleL 184
321 ATTTTTCANANTACCAAGCAGCCGACCCACCCCAATGCTTATCT 370
184 euArgLeuLeuArgTyrIleValArgLeuValTrpArgMetHis 198
371 TACGACTGTTACGTTACATGTGCGCGCTGGTGTGGAGAATGCAT 414

```

```

seq_name: gb_un:AX031311
seq_documentation_block:
LOCUS AX031311 416 bp DNA linear UNA 20-SEP-2000
DEFINITION Sequence 7 from Patent WO9914321.
ACCESSION AX031311
VERSION AX031311.1 GI:10278639
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE
1 (bases 1 to 416)
O'Reilly,L., Ruthalakath,H., Adams,J., O'Connor,L., Cory,S.,
Huang,D.C. and Strasser,A.
Novel therapeutic molecules
Patent: WO 9914321-A 25-MAR-1999;
INST MEDICAL W & E HALL (AU); PUTHALAKATH HANSA (AU); REILLY
LORRAINE O (AU); ADAMS JERRY (AU); CONNOR LIAM O (AU); CORY
SUZANNE (AU); HUANG DAVID C S (AU); STRASSER ANDREAS (AU)
FEATURES
Location/Qualifiers
1. 416
/organism="unidentified"
/db_xref="taxon:32644"
1. >414
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09661.1"
/db_xref="GI:10278640"
/translation="MAKQPSDVSSECDREGROLOPAERPPOLRPGAPTSIQTEPODORS
PAPMSCDKSTQTPSPPCQAFNHVLSAMASMRQAEADMRPEIWIQAQLRIGDEFNAY
YARRVFLNNTQAEEDHPRMVILRLRYIVRLVRMH"
BASE COUNT 113 a 113 c 103 g 87 t
ORIGIN
alignment_scores:
  Quality: 696.00      Length: 198
  Ratio: 5.043        Gaps: 1
  Percent Similarity: 69.697   Percent Identity: 69.697
alignment_block:
US-09-508-832-10 x AX031311 ..
Align seg 1/1 to: AX031311 from: 1 to: 416
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
1 ATGGCAAGGCAACCTCTGATGTAAGTCTGAGTGTGACCGAGAAGGTAG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProt 34
51 ACAATTGCAGCTCGCGAGAGGCGCTCCACGCTCAGACCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
101 CCTCCTACAGACAGAGCCACAA..... 123
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
123 ..... 123
67 oAlaSerProGlyPropheAlaThrArgSerProLeuPheIlePheMeta 84
123 ..... 123
84 rArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
123 ..... 123
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
124 ...GACAGGAGCCGACGCCATGAGTTGTGACAAATCAACACAAACCCC 170

```

117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
 118 |||||
 171 AAGTCTCTTGGCCAGCCCTTCAACACATATCTAGTGCATGCTTCCA 220
 134 eTarGlnAlaGluProAlaAspMetArgProGluIleTrrPileAlaGln 150
 221 TGAGGAGGCTGAACCTTGCAGATATGCGCCAGAGATATGATGCGCCAA 270
 151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgVa 167
 271 GAGTTGCGCGCTATCGGAGCAGAGTTTAAAGCTTACTATGCAAGGAGGT 320
 167 lPheLeuAsnTyrGlnAlaAlaGluAspHisProArgMetValIleL 184
 321 ATTTTGAATATATCAACAGCAGCCGAGACCCACCGATGGTTATCT 370
 184 euArgLeuLeuArgTyrIleValArgLeuValTrrArgMethHis 198
 371 TAGGACTGTACGTATACATTTGTCGCCCTGGTGTGAGATGCAT 414

seq_name: gb_pr:AF032458

seq documentation_block:
 LOCUS AF032458 417 bp mRNA linear PRI 19-FEB-1998
 DEFINITION Homo sapiens BimL mRNA, complete cds.
 ACCESSION AF032458
 VERSION AF032458.1 GI:2895497
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 417)
 O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
 Cory,S. and Huang,D.C.
 Bim: a novel member of the Bcl-2 family that promotes apoptosis
 TITLE
 JOURNAL ENBO J. 17 (2), 384-395 (1998)
 MEDLINE 96094360
 PUBMED 9430630
 REFERENCE
 2 (bases 1 to 417)
 O'Connor,L., Strasser,A., O'Reilly,L.A., Hausmann,G., Adams,J.M.,
 Cory,S. and Huang,D.C.S.
 Direct Submission
 TITLE
 JOURNAL Submitted (03-NOV-1997) Molecular Genetics of Cancer, The Walter
 Eliza Hall Institute of Medical Research, PO Royal Melbourne
 Hospital, Parkville, Victoria 3050, Australia
 FEATURES
 source
 1..417
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..417
 /note="pro-apoptotic BH3-containing Bcl-2 family member"
 /codon_start=1
 /product="BimL"
 /protein_id="AAC39594.1"
 /db_xref="GI:2895498"
 /translation="MAKQPSDVSSECDREGQLQPAERPPQLRPGAPTSLQTEPDQRS
 PAPMCDSTQTPSPCFNHLUSANASRQAEAPMRPELWIAQELRRIGDEFNAY
 YARRVELNNYQAEHDHPWVILRLIYVLRVWRMH"

BASE COUNT 114 a 113 c 103 g 87 t
 ORIGIN

alignment_scores:
 Quality: 696.00 Length: 198
 Ratio: 5.043 Gaps: 1
 Percent Similarity: 69.697 Percent Identity: 69.697

alignment_block:
 us-09-508-832-10 x AF032458 ..

Align seg 1/1 to: AF032458 from: 1 to: 417

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
 18 |||||
 1 ATGGCAACCAACCTTCTGATGTAAAGTTCTGAGTGTGACCGAGAAGGTAG 50
 17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
 51 ACAATTTGCAGCCTGCGAGAGGCGCTCCCGAGCTCAGACCTGGGGGCCCTA 100
 34 hrSerLeuGlnThrGluProGlnGlnAsnProGluGlyAsnHisGlyGly 50
 101 CCTCCCTACAGACAGAGCCAA..... 123
 51 GluGlyAspSerCysProHisGlySerProGlnGlnProLeuAlaProPr 67
 123 123
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
 123 123
 84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
 123 123
 101 ThrAspArgSerProAlaProMetSerCysAspIysSerThrGlnThrPr 117
 124 ...GACAGGAGGCCAGCACCCATGAGTGTGACAAATCAACACAAACCC 170
 117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
 171 AAGTCTCTTGGCCAGGCTTCAACCACTATCTCAGTGCATGCTTCCA 220
 134 eTarGlnAlaGluProAlaAspMetArgProGluIleTrrPileAlaGln 150
 221 TGAGGAGGCTGAACCTTGCAGATATGCGCCAGAGATATGATGCGCCAA 270
 151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgVa 167
 271 GAGTTGCGCGCTATCGGAGCAGAGTTTAAAGCTTACTATGCAAGGAGGT 320
 167 lPheLeuAsnTyrGlnAlaAlaGluAspHisProArgMetValIleL 184
 321 ATTTTGAATATATCAACAGCAGCCGAGACCCACCGATGGTTATCT 370
 184 euArgLeuLeuArgTyrIleValArgLeuValTrrArgMethHis 198
 371 TAGGACTGTACGTATACATTTGTCGCCCTGGTGTGAGATGCAT 414

CC expression of Bim activity is useful in regulating inhibition or
 CC prevention of cell death or degeneration such as under cytotoxic
 CC conditions during e.g. gamma-irradiation and chemotherapy or during
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
 CC hypoxia, degenerative diseases or for prolonging the survival of
 CC cells being transplanted for treatment of disease. Since Bim is
 CC expressed in germ cells, modulating Bim expression or Bim activity
 CC is useful, e.g. as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.

XX Sequence 596 BP: 145 A: 175 C: 146 G: 130 T: 0 other:

alignment_scores:
 Quality: 1065.00 Length: 198
 Ratio: 5.379 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-508-832-10 x AAX24997

Align seg 1/1 to: AAX24997 from: 1 to: 596

1 MetAlaLysGlnProSerAspValSerGluCysAspArgGluGlyAr 17
 1 ATGCGAAGCAACCTCTGATGTAAGTCTGAGTGTGACCGAGAGGTAG 50
 17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
 51 ACAATTCACGCTCGGAGAGGCGCTCCCGAGCTCAGACCTGGGGCCCTA 100
 34 hrSerLeuGlnThrGluProGlnGlnGlnGlnGlnGlnGlnGlnGly 50
 101 CCTCCTTACACAGACAGACGACACAGCTAATCTCGAAGGCAATCAGGAGGT 150
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaP:opr 67
 151 GAAGGGGACAGCTGCCCGACGCGAGCCCTCAGGGCGCGCTGGCCGCCACC 200
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
 201 TGGCAGGCGCTGGCGCTTTGCTTACAGATCCCGCTTTTCATCTTATGA 250
 84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
 251 GAAGATCCTCCCTGCTCTCGATCCTCCAGTGGGTATTTCTCTTTTGAC 300
 101 ThrAspArgSerProAlaProMetSerCysAspIysSerThrGlnThrPr 117
 301 ACAGACAGGAGCCACCCACCCATGAGTTGTGACAAATCAACACACCC 350
 117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerH 134
 351 AAGTCTCTCTTGGCAGGCGCTTCAACCACTATCTCAGTGCATGGCTTCCA 400
 134 eTArgGlnAlaGluProAlaAspMetArgProGluIleTTrpIleAlaGln 150
 401 TGAGGCAAGCTGAACCTGAGATATCCCGCCAGACATATGATCGCCCAA 450
 151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgVa 167
 451 GAGTTCGGCGGTATCGGAGACGAGTTTAACCGCTTACTATCAAGGAGGCT 500
 167 lPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetValIleL 184
 501 ATTTTGAATAATACCAAGCAGCGGAGACCCACCCAGGAATGTTATCT 550
 184 euArgLeuLeuArgTyrIleValArgLeuValTPrArgMetHis 198
 551 TAGCACTGTACCTACATGTGCCCGCTGGTGGAGAAATGCAT 594

seq_name: /SIDS1/gc9data/geneseq/geneseqn-emb1/NA1999.DAT: AAX24995

seq_documentation_block:

ID AAX24995 standard; cDNA; 590 BP.

XX AAX24995:

DT 05-JUL-1999 (first entry)

DE Murine Bcl-2 interacting mediator of cell death Bim-EL cDNA.

XX Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; mouse; cancer; autoimmune disease;
 KW degenerative disease; therapy; contraceptive; splice variant;
 KW isoform; ss.

XX Mus musculus.

XX W09914321-A1.

XX 25-MAR-1999.

XX 17-SEP-1998: 98MO-AU00772.

XX 24-SEP-1997: 97AU-0009373.

XX 17-SEP-1997: 97AU-0009263.

XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;

XX WPI: 1999-244030/20.

XX P-PSDB: AAW98156.

XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 PT treatment

XX Claim 3: Page 96-97: 145pp: English.

CC The present sequence encodes the extra long form (EL) of murine Bim,
 CC or Bcl-2 interacting mediator of cell death (see AAW98156), a novel
 CC member of the Bcl-2 family that is capable of inducing cell death
 CC (apoptosis) and which acts as a 'death-ligand' for certain members
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
 CC only Bcl-2 homology region which it encompasses is BH3. It is the
 CC result in the expression of a variety of isoforms, i.e. Bim-S,
 CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim
 CC isoforms were obtained from a T lymphoma cDNA library using human
 CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to
 CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have
 CC also been identified (see AAW98157-59). Binding the dynein light
 CC chain was shown to regulate the pro-apoptotic activity of Bim.
 CC Bim-S, the splice variant which does not bind to dynein light
 CC chain, is a much more potent killer than either Bim-L or Bim-EL.
 CC The invention provides variants (see AAW98159-68) of murine and human
 CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate
 CC with a dynein light chain. The identification of Bim permits the
 CC identification and rational design of a range of products for use
 CC in therapy, diagnosis, antibody generation and involving modulation
 CC of physiological cell death. These therapeutic molecules may act
 CC as either antagonists or agonists of Bim's function and will be
 CC useful in cancer, autoimmune or degenerative disease therapy.
 CC Increased Bim expression or Bim activity is useful, e.g. for
 CC treatment of prophyllaxis in conditions such as cancer and deletion
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
 CC expression of Bim activity is useful in regulating inhibition or
 CC prevention of cell death or degeneration such as under cytotoxic
 CC conditions during e.g. gamma-irradiation and chemotherapy or during
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
 CC hypoxia, degenerative diseases or for prolonging the survival of
 CC cells being transplanted for treatment of disease. Since Bim is
 CC expressed in germ cells, modulating Bim expression or Bim activity
 CC is useful, e.g. as a contraceptive or method of sterilization by

CC preventing generation of fertile sperm.
XX
SQ Sequence 590 BP; 137 A; 178 C; 150 G; 125 T; 0 other;

alignment_scores:			
Quality:	908.00	Length:	200
Ratio:	4.935	Gaps:	2
Percent Similarity:	92.000	Percent Identity:	86.500

alignment_block:

US-09-508-832-10 x AAX24995

Align seq 1/1 to: AAX24995 from: 1 to: 590

```

1  MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
1  ATGGCCACGACCAACCTTCTGATGTAAAGTTCTGAGTGTGACAGAGATGG 50
17  gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
51  ACAATTGCAGCTGCTCAGAGGCTCCCCAGCTCAGGCTGGGGCCCTTA 100
34  hSertLeuGlnThrGluProGlnGlyAsnProGluGlyGlyGly 50
101  CCTCCCTACAGACAGAACCGAAGGTAAATCCCGAC.....GGC 138
51  GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProP 67
139  GAAGGGACCGCTGCCCCCAACGGCAGCCCTCAGGSCCGCTGGCGCCACC 188
67  oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
189  GGCCAGCCCTGGCCCTTTTCTACACAGATCCCCACTTTTCATCTTTGTGA 238
84  rArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
239  GAAGATCTTCTGCTGTCCCGCTCTCCAGTGGGTATTTCTCTTTTGAC 288
101  ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
289  ACAGACAGAGCGCGGACCCATGACTGTGACAAGTCAACACAAACCCCC 338
117  oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
339  AAGTCTCTCTTGGCAGGCCCTCAACCACATATCTCAGTGCANTGGCTTCCA 388
134  eArgGlnAla.....GluProAlaAspMetArgProGluIleTrpIle 148
389  TAGCAGACGTCTCAGGAGGAACCTTGAAGATCTCGGCCCGGAGATACGGATT 438
149  AlaGlnGluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaAr 165
439  GCACAGAGAGCTGCGCGGATCGGAGCAGGATTCACAGAAACTTACACAAG 488
165  rArgValPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetV 182
489  GAGGTGTTTGCAAATGATTAATCCGCGAGGCTGAGACACCCCTCAANTGG 538
182  alIleLeuArgLeuLeuArgTyrIleValArgLeuValTrpArgMetHis 198
539  TTATCTTACAACTGTACGCTTTATCTTCCTCGCTCGGTATGAGGAAGCAT 588
seq_name: /SIDS1/gcgdata/qnensed/qneseq/qneseq-emb1/NA1999..DAT..AAAX24

```

Bim-L; Bcl-2 interacting mediator of cell death; apoptosis; cell cycle; human; cancer; autoimmune disease; degenerative disease; therapy; contraceptive; splice variant; isoform; ss.

Homo sapiens.

WO914321-A1.

25-MAR-1999.

17-SEP-1998: 98WO-AU00772.

24-SEP-1997: 97AU-0009373.

17-SEP-1997: 97AU-0009263.

(HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L; Puthalakath H, Strasser A;

WPI; 1999-244030/20.

P-PDB; AAW98157.

New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment

Claim 7; Page 99-100; 145pp; English.

The present sequence encodes the long form (L) of human Bim, or Bcl-2 interacting mediator of cell death (see AAW98157), a novel member of the Bcl-2 family that is capable of inducing cell death (apoptosis) and which acts as a 'death-ligand' for certain members of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the only Bcl-2 homology region which it encompasses is BH3. It is the only BH3-only protein for which splice variants exist. These result in the expression of a variety of isoforms, i.e. Bim-S, Bim-L and Bim-EL. cDNAs encoding human Bim-L and Bim-EL (see AAW98158) were isolated from embryo and liver cDNA libraries using mouse Bim cDNA. Murine Bim-S, Bim-L and Bim-EL isoforms (see AAW98154-56) are also provided. The human Bim gene maps to chromosome 2 at bands 2q12-2q13. Binding the dynein light chain was shown to regulate the pro-apoptotic activity of Bim. Bim-S, the splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain. The identification of Bim permits the identification and rational design of a range of products for use in therapy, diagnosis, antibody generation and involving modulation of physiological cell death. These therapeutic molecules may act as either antagonists or agonists of Bim's function and will be useful in cancer, autoimmune or degenerative disease therapy. Increased Bim expression or Bim activity is useful, e.g. for treatment of prophylaxis in conditions such as cancer and deletion of autoreactive lymphocytes in autoimmune disease. Decreased Bim expression of Bim activity is useful in regulating inhibition or prevention of cell death or degeneration such as under cytotoxic conditions during e.g. gamma-irradiation and chemotherapy or during HIV/AIDS or other viral infections, ischemia, myocardial infarction, hypoxia, degenerative diseases or for prolonging the survival of cells being transplanted for treatment of disease. Since Bim is expressed in germ cells, modulating Bim expression or Bim activity is useful, e.g. as a contraceptive or method of sterilization by preventing generation of fertile sperm.

Sequence 416 BP; 113 A; 113 C; 103 G; 87 T; 0 other;

Percent Similarity: 69.697 Percent Identity: 69.697
 alignment_block:
 US-09-508-832-10 x AAX24996
 Align seg 1/1 to: AAX24996 from: 1 to: 416

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
 1 AYGGCAAGCAACCTCTGATGTAAGTCTGAGTGTGACCGAGAAGTAG 50
 17 gClnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
 51 ACAATTGCAGCTCGCGAGAGGCTCCCGAGCTCAGACCTGGGCGCCCTA 100
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
 101 CTTCCCTACAGACAGCCCAAA..... 123
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
 123 123
 67 oAlaSerProGlyPropheAlaThrArgSerProLeuPheIlePheMetA 84
 123 123
 84 rGArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
 123 123
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
 124 ...GACAGGAGCCGACGACCCATGAGTGTGACAAATCAACACAAACCC 170
 117 oSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
 171 NAGTCCCTTTCGAGCGCTTCAACCACTATCTCAGTGCATGCGTTCCA 220
 134 eArgGlnGluProAlaAspMetArgProGluIleThrPheAlaGln 150
 221 TCAGCGAGCTGAAACCTGCAAGATATCGCCGAGATATGATGATGCGCCAA 270
 151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaArgArgVa 167
 271 CAGTTGGCGGCTATCGGAGACGAGTTTAAACGCTTACTATGCAAGCAGGT 320
 167 lPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetValIleL 184
 321 ATTTTGTGATAATACACAGCAGCCGACGACCCAGCAATGTTATCT 370
 184 eArgLeuLeuArgTyrIleValArgLeuValTrpArgMethIs 198
 371 TACGACTGTTACGTTACATTTGTCGCTGCGTGTGAGAAATGCAT 414

seq_name: /SDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: AAX24994
 seq_documentation_block:
 ID_AAX24994 standard; cDNA; 422 BP.
 XX AC AAX24994;
 XX DT 05-JUL-1999 (first entry)
 XX DE Murine Bcl-2 interacting mediator of cell death Bim-L cDNA.
 XX KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; mouse; cancer; autoimmune disease;
 KW degenerative disease; therapy; contraceptive; splice variant;
 KW isoform; ss.
 XX OS Mus musculus.
 XX

PN WO9914321-A1.
 XX 25-MAR-1999.
 XX 17-SEP-1998; 98WO-AU00772.
 XX 24-SEP-1997; 97AU-0009373.
 PR 17-SEP-1997; 97AU-0009263.
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI puthalakath H, Strasser A;
 XX MPI: 1999-244030/20.
 DR P-PSDB; AAW98155.
 XX
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 treatment
 PS Claim 3: Page 94-95; 145pp; English.
 CC The present sequence encodes the long form (L) of murine Bim, or
 CC Bcl-2 interacting mediator of cell death (see AAW98155), a novel
 CC member of the Bcl-2 family that is capable of inducing cell death
 CC (apoptosis) and which acts as a 'death-ligand' for certain members
 CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
 CC only Bcl-2 homology region which it encompasses is BH3. It is the
 CC only BH3-only protein for which splice variants exist. These
 CC result in the expression of a variety of isoforms, i.e. Bim-S,
 CC Bim-L and Bim-EL (see AAW98154-56). cDNAs encoding these murine Bim
 CC isoforms were obtained from a T lymphoma cDNA library using human
 CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to
 CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have
 CC also been identified (see AAW98157-58). Binding the dynein light
 CC chain was shown to regulate the pro-apoptotic activity of Bim.
 CC Bim-S, the splice variant which does not bind to dynein light
 CC chain, is a much more potent killer than either Bim-L or Bim-EL.
 CC The invention provides variants (see AAW98159-68) of murine and human
 CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate
 CC with a dynein light chain. The identification of Bim permits the
 CC identification and rational design of a range of products for use
 CC in therapy, diagnosis, antibody generation and involving modulation
 CC of physiological cell death. These therapeutic molecules may act
 CC as physical antagonists or agonists of Bim's function and will be
 CC useful in cancer, autoimmune or degenerative disease therapy.
 CC Increased Bim expression or Bim activity is useful, e.g. for
 CC treatment or prophylaxis in conditions such as cancer and deletion
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
 CC expression of Bim activity is useful in regulating inhibition or
 CC prevention of cell death or degeneration such as under cytotoxic
 CC conditions during e.g. gamma-irradiation and chemotherapy or during
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
 CC hypoxia, degenerative diseases or for prolonging the survival of
 CC cells being transplanted for treatment of disease. Since Bim is
 CC expressed in germ cells, modulating Bim expression or Bim activity
 CC is useful, e.g. as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 XX SQ Sequence 422 BP; 112 A; 116 C; 109 G; 85 T; 0 other;

alignment_scores:
 Quality: 589.00 Length: 200
 Ratio: 4.566 Gaps: 2
 Percent Similarity: 64.500 Percent Identity: 60.000
 alignment_block:
 US-09-508-832-10 x AAX24994
 Align seg 1/1 to: AAX24994 from: 1 to: 422

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17

|||||
1 ATGGCCAGCAACCTTCTGATGTAGTTCTGACGTGTGACAGAGAGGTGG 50
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
51 ACAATTCAGGCTCTCTGAGAGGCTCCCGAGCTCAGGCTGGGGCCCTA 100
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
101 CCTCCTACAGACAGAACCGCAA..... 123
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
123 123
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
123 123
84 rArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
123 123
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
124 ...GACAGAGCCCGGACCATGAGTTGTGACAGTCAACACAAACCC 170
117 oSerProCysGlnAlaPheAsnHisLysLeuSerAlaMetAlaSerM 134
171 AAGTCCTCTGCGGAGCCTTCAACCACTATCTCAGTGCATGCAATGCC 220
134 eTargGlnAla..... GluProAlaAspMetArgProGluIleTrpIle 148
221 TAGCAGCTCTCAGAGGACCTGAGTCTGCGCCCGGAGATACGGATT 270
149 AlaGlnGluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyrAlaAr 165
271 GCACAGGAGCTGCGCGGATCGGAGACGAGTTCAACGNAACCTTACACAAG 320
165 gArgValPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMetV 182
321 GAGGCTGTTTGCAAATGATTACCGCGAGGCTGAAGACCAACCCCTCAA 370
182 alrLeuArgLeuLeuArgTyrIleValArgLeuValTrpArgMetHis 198
371 TTATCTTACACTGTATACGCTTTATCTCTCGTCTGTATGGAGAGGCAT 420

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: AAS65350

seq_documentation_block:

ID AAS65350 standard; cDNA; 442 BP.

XX
AC AAS65350;

XX
DT 13-FEB-2002 (first entry)

XX
DE DNA encoding novel human diagnostic protein #1154.

XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX
OS Homo sapiens.

XX
PN WO200175067-A2.

XX
PD 11-OCT-2001.

XX
PF 30-MAR-2001; 2001WO-US08631.

XX
PR 31-MAR-2000; 2000US-0540217.

XX
PR 23-AUG-2000; 2000US-0649167.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB: ABG01163.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1: SEQ ID No 1154; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 442 BP; 94 A; 154 C; 105 G; 89 T; 0 other;

alignment_scores:
Quality: 497.00 Length: 92
Ratio: 5.402 Gaps: 0
Percent similarity: 100.000 Percent identity: 100.000

alignment_block:

US-09-508-832-10 x AAS65350 ..

Align seg 1/1 to: AAS65350 from: 1 to: 442

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
167 ATGGCAAGCAACCTTCTGATGTAGTTCTGACGTGTGACAGAGAGGTAG 216
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
217 ACAATTCAGGCTCTCTGAGAGGCTCCCGAGCTCAGGCTGGGGCCCTA 266
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
267 CCTCCTACAGACAGACAGCCACAGGTAATCTCTGAGGCAATCAGCGAGGT 316
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
317 GAAGGGGACAGCTGCCCCACGCGAGCCCTCAGGGCGCGCTGGCCACC 366
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
367 TGCCAGCCCTGGGCCCTTTGCTACCCAGATCCCGGCTTTTCATCTTTATGA 416
84 rArgSerSerLeuLeuSerArgSer 92
417 GAAGATCCTCCTGCTGCTCGCATCC 442

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: AAX24993

CC is useful, e.g. as a contraceptive or method of sterilization by
CC preventing generation of fertile sperm.

ID AAX24993 standard; CDNA: 332 BP.

XX
AC AAX24993:

DT 05-JUL-1999 (first entry)

DE Murine Bcl-2 interacting mediator of cell death Bim-S cDNA.

KW Bim-S; Bcl-2 interacting mediator of cell death; apoptosis;

degenerative disease; therapy; contraceptive; splice variant;

XX
XX

XX

XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX 25 100 2

PR 17-SEP-1997; 97AU-0009263.

PA (HALL-) HALL INST MEDICAL RE
YY

PI Adams J, Cory S, Huang DCS
PI Ruthelekathu Stressor A:

XX
DR
WPI: 1999-244030/20

DR P-PSDB; AAW98134.
XX

PT	NEW ISOLATED member of the treatment
PT	treatment

xx PS Claim 3: page 92: 14

The present sequence

CC member of the Bcl-2 family that is capable of inducing cell death

CC of 'the pro

only BH3-only protein for which s

Blim-L and Blm-EL (see AW981354-56). cDNAs encoding these murine Blm isoforms were obtained from a T lymphoma cDNA library using human recombinant Bcl-2 protein. The murine Blm gene has been mapped to chromosome 2 at bands F3-6. Human Blm-L and Blm-EL isoforms have also been identified (see AW981357-58). Binding the dynein light chain was shown to regulate the pro-apoptotic activity of Blm. Blm-S, the splice variant which does not bind to dynein light chain, is a much more potent killer than either Blm-L or Blm-EL. The invention provides variants (see AW98159-68) of murine and human Blm-L or Blm-EL that cannot bind, couple or otherwise associate with a dynein light chain. The identification of Blm permits the identification and rational design of a range of products for use in therapy, diagnosis, antibody generation and involving modulation of physiological cell death. These therapeutic molecules may act as either antagonists or agonists of Blm's function and will be useful in cancer, autoimmune or degenerative disease therapy. Increased Blm expression or Blm activity is useful, e.g. for treatment or prophylaxis in conditions such as cancer and deletion of autoreactive lymphocytes in autoimmune disease. Decreased Blm expression of Blm activity is useful in regulating inhibition or prevention of cell death or degeneration such as under cytotoxic conditions during e.g. gamma-irradiation and chemotherapy or during HIV/AIDS or other viral infections, ischemia, myocardial infarction, hypoxia, degenerative diseases or for prolonging the survival of cells being transplanted for treatment of disease. Since Blm is expressed in germ cells, modulating Blm expression or Blm activity

```

DE Human polynucleotide SEQ ID NO 9344.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; Cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB: AA009353.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1: SEQ ID NO 9344; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 371 BP; 106 A; 95 C; 81 G; 89 T; 0 other;

alignment_scores:
    Quality: 123.00      Length: 23
    Ratio: 5.348         Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-508-832-10 x AAI89284/rev ..

Align seg 1/1 to reverse of: AAI89284 from: 1 to: 371

176 GluAspHisProArgMetValIleLeuArgLeuArgLeuArgTyrIleValAr 192
|||||
371 GAAGACCAACCACCAAGTGTATCTACGACTGTACGTTACATGTCGC 322
|||||

192 gLeuValTrpArgMethHis 198
|||||
321 CTGGTGTGGAGATGCAT 303

seq_name: /SIDSL/ycgdata/geneseq/geneseq-emb1/NAI998.DAT:AAV35620

seq_documentation_block:
ID AAV35620 standard; DNA; 32367 BP.
XX
XX AAV35620;

```

```

XX
XX 07-SEP-1998 (first entry)
XX
XX Human SHOX (short stature homeobox containing gene) gene sequence.
XX
XX Homeobox domain; human growth gene; growth regulation; growth defect;
KW Turner's syndrome; short stature homeobox containing gene; SHOXa;
KW SHOX; bone disease; osteoporosis; calcium regulation; short stature;
KW transcription factor A; ss.
XX
XX Homo sapiens.
XX
XX WO9814568-A1.
XX
XX 09-APR-1998.
XX
XX 29-SEP-1997; 97MO-EF05355.
XX
XX 16-JAN-1997; 97EP-0100583.
XX
XX 01-OCT-1996; 96US-0027633.
XX
XX (RAPP/) RAPPOLD-HOERBRAND G.
XX
XX Rao E, Rappold-hoerbrand G;
XX
XX WPI; 1998-271719/24.
XX
XX New human growth genes - used to develop products for the diagnosis
PT and treatment of human growth defects such as short stature, e.g.
PT Turner's syndrome
XX
XX Claim 19; Pages 51-67; 84pp; English.
XX
XX This is the human SHOX gene sequence containing the PAR1 region. The
CC gene region corresponding to short stature has been identified as a
CC region of approximately 500 kb in the PAR1 region of the X and Y
CC chromosomes. Three genes in this region have been identified as
CC candidates for the short stature gene. These genes were designated SHOX
CC (also referred to as SHOX93 or HOX93), PET92 and SHOT (SHOX-like homeobox
CC gene on chromosome three). The SHOX gene has two separate splicing sites
CC resulting in two variations SHOXa and SHOXb. The specification provides
CC sequences of SHOX (short stature homeobox-containing) genes SHOX ET92,
CC SHOXa, SHOXb, SHOT and exons of the SHOX genes as shown in AAV35610 to
CC AAV35621 and protein sequences of the human growth protein transcription
CC factor SHOXa, SHOXb and SHOT as shown AAW60573 to AAW60575. The novel
CC genes are responsible for human growth. Defects in the genes can cause
CC short stature, e.g. Turner's syndrome. The products can be used to
CC develop agents for the treatment of short stature or other human growth
CC disorders. The products can also be used for providing a mitogenic effect
CC on cells, e.g. for the treatment of bone diseases such as osteoporosis
CC and diseases involved with disturbance in the bone calcium regulation.
XX
XX Sequence 32367 BP; 7627 A; 8130 C; 8564 G; 8043 T; 3 other;

alignment_scores:
    Quality: 119.50      Length: 177
    Ratio: 1.271         Gaps: 13
    Percent Smilarity: 53.107    Percent Identity: 33.898

alignment_block:
US-09-508-832-10 x AAV35620/rev ..

Align seg 1/1 to reverse of: AAV35620 from: 1 to: 32367

9 SerSerGluCysAspArgGluGlyArg.....GlnLeuGlnProAlaG1 23
|||||
14144 TCCCCAAAGCCTCTCCCGGAGACACACGAGTGTCCAAAGCCTCTCCCGGGTCCAC 14095
|||||

23 uArgProProGlnLeu.....ArgProGlyAlaProThrSerLeug 37
|||||
14094 CAGGTCCCAAGGCTCTCTCCCGAGACACACGAGTGTCCCGCCAGCCTCTCCGC 14045
|||||

```


109 rCysAspLysSerThrGlnThrPro...SerProProCysGlnAlaPheA 125
 453 TCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 404
 125 snHsTyrLeuSerAlaMetAlaSerMetArgGlnAlaGluProAla 140
 403 CT 357

seq_name: /SDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: AAD07132

seq_documentation_block:

ID AAD07132 standard; CDNA; 2803 BP.

XX AC AAD07132;

XX DT 06-AUG-2001 (first entry)

XX DE Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #2 CDNA.
 XX KW Dog; X-linked progressive retinal atrophy 2; XLPRA2; genetic marker;
 KW retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
 XX Miniature Schnauzer; mutant; mutelin; ss.
 XX OS Canis familiaris.
 XX OS Synthetic.

XX FH Key Location/Qualifiers
 XX FT CDS 1..1149
 FT /*tag= a
 FT /product= "Canine retinitis pigmentosa GTPase
 FT regulator mutant"
 FT /note= "CDS does not include start codon"
 FT /partial
 FT mutation
 FT replace (931..932, ACAG)
 FT /tag= b
 FT /note= "This deletion results in the change of
 FT amino acids and ends in a premature stop codon."

XX PN W0200138578-AL.

XX PD 31-MAY-2001.

XX PF 21-NOV-2000; 2000MO-US31940.

XX PR 24-NOV-1999; 99US-0167365.

XX PA (CORR) CORNELL RES FOUND INC.

XX PI Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;

XX DR WPI; 2001-367707/38.

XX DR P-PSDB; AAE02399.

XX PT Identifying dogs with or carrying X-linked progressive retinal atrophy
 XX PT by detecting retinitis pigmentosa GTPase regulator gene mutation,
 XX PT useful when breeding Husky, Samoyed and Miniature Schnauzer

XX PS Claim 75: Page 34-35; 86pp; English.

XX CC The invention relates to a method for identifying dogs which are
 CC genetically normal, are carriers of, or are affected with X-linked
 CC progressive retinal atrophy (XLPRA), by testing a biological sample with
 CC genetic markers that co-segregate with a XLPRA gene locus. The invention
 CC also relates to canine retinitis pigmentosa GTPase regulator (RPGR). RPGR
 CC mutants and their corresponding nucleic acid molecules. The mutated RPGR
 CC genes are responsible for the XLPRA in dogs. Methods are used to select
 CC dogs for breeding so that dogs carrying the mutated locus are eliminated
 CC from the breeding stock. The method particularly applies to Siberian
 CC Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
 CC the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
 CC and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
 CC The present cDNA sequence is the mutant open reading frame (ORF) 15
 CC encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant

CC found in XLPRA2-affected dogs. This mutant is obtained by deleting
 CC 'GA' nucleotide bases from position 932 to 933 of the wild-type
 CC canine RPGR cDNA.

XX S0 Sequence 2803 BP; 1028 A; 287 C; 939 G; 549 T; 0 other;

alignment_scores:

Quality: 113.00 Length: 149

Ratio: 1.314 Gaps: 6

Percent Similarity: 57.718 Percent Identity: 27.517

alignment_block:

US-09-508-832-10 x AAD07132/rev ..

Align seg 1/1 to reverse of: AAD07132 from: 1 to: 2803

5 ProSerAspValSerSerGluCysAspArgGluGlyArgGlnLeuGlnPr 21
 770 CTTCTTCT 721
 21 oAlaGluArgProGlnLeuAlaProGlyAlaProThrSerLeuGln 38
 720 CTCCT 671
 38 hrGluProGlnGlyAsnProGluGlyAsnHisGlyGlyGluGlyAspSer 54
 670 CTCCT 630
 55 CysProHisGlySerPro.....GlnGlyProLeuAlaProAla... 68
 629 TCCCT 580
 69SerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
 579 TAATCT 546
 84 rGArgSerSerLeuSerArgSerSerGlyTyrPheSerPheAsp 100
 545TCT 504
 101 ThrAspArgSer.....ProAlaProMetSe 109
 503 GCT 454
 109 rCysAspLysSerThrGlnThrPro...SerProProCysGlnAlaPheA 125
 453 TCCT 404
 125 snHsTyrLeuSerAlaMetAlaSerMetArgGlnAlaGluProAla 140
 403 CT 357

seq_name: /SDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: AAD07130

seq_documentation_block:

ID AAD07130 standard; cDNA; 2805 BP.

XX AC AAD07130;

XX DT 06-AUG-2001 (first entry)

XX DE Canine retinitis pigmentosa GTPase regulator (RPGR) CDNA.

XX KW Dog; X-linked progressive retinal atrophy; XLPRA; genetic marker;
 KW retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
 XX Miniature Schnauzer; ss.

XX OS Canis familiaris.

XX FH Key Location/Qualifiers

XX FT CDS 1..1569

XX FT /*tag= a


```
alignment_block:
US-09-508-832-10 x AAL24943/rev ..
Align seg 1/1 to reverse of: AAL24943 from: 1 to: 766
21 ProAlaGluArgProProGlnLeuArgProGlyAlaProThrSerLeuG1 37
||||| : : : : : ||||| : : : : :
644 CTGGCTCAAAAACCCCTCCAGGGCGCGGGTGTCTCCTCTACC..... 600
37 nThrGluProGlnGlyAsnProGlyAla.....HisGlyGlyGluG 52
||||| : : : : : ||||| : : : : :
599 .....CAGGGGAATCCTCAGGGGATATTAAGATGGGGCCACAG 560
52 lAspSer.....CysPro 56
||||| : : : : : ||||| : : : : :
559 GGGAAAGGAAAGCGCGCGTAAATCTTACATACCCCGGAGCTTTGTGAT 510
57 HisGly .....SerProGlnGlyPro..... 63
||||| : : : : : ||||| : : : : :
509 TCGGGTCCCCAAACCGCGCCAGTCCCAAGCGCGGAAACCCACATTC 460
64 .....LeuAlaProAlaSerProGlyProPheAlaThrArgS 77
: : : : : ||||| : : : : :
459 CCATCCAGAGCGATGCAAAACCCAGACCCAGTCCCATTCACCTACCAC 410
77 erProLeuPheIlePheMetArgSerSerLeuLeuSerArgSer 93
||||| : : : : : ||||| : : : : :
409 CCCC.....CATTCAGACATACCCAAACATAGACAG 378
94 SerGlyTyrPheSerPheAspThrArgSerProAlaProMetSerCy 110
: : : : : ||||| : : : : :
377 ACAGGCAAA.....GACACAGACCCAGACCCACCATC..... 342
110 sAspLysSerThrGlnThrProSerProProCysGlnAlaPheAsnH127
||||| : : : : : ||||| : : : : :
341 .....ACACCCAGACCC.....ANTCAT. 324
127 yrLeuSerAlaMetAlaSerMetArgGlnAlaGluProAlaAspMetArg 143
: : : : : ||||| : : : : :
323 .....CCAGGCCAAACCCAAACACCCAGCGGAGAACCCACCCAGCAGA 279
144 ...ProGlu 145
|||||
278 AACCCAGAG 270
seq_name: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS84424
seq_documentation_block:
ID AAS84424 standard; cDNA; 2370 BP.
XX AC AAS84424;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #20228.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX OS food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX PN Homo sapiens.
XX PD W0200175067-A2.
XX PF 11-OCT-2001.
XX PR 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX
```

```
BI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG20237.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 20228; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2370 BP; 474 A; 623 C; 753 G; 520 T; 0 other;

alignment_scores:
Quality: 112.00 Length: 129
Ratio: 1.836 Gaps: 7
Percent Similarity: 47.287 Percent Identity: 31.008

alignment_block:
US-09-508-832-10 x AAS84424/rev ..
Align seg 1/1 to reverse of: AAS84424 from: 1 to: 2370
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
: : : : : ||||| : : : : : ||||| : : : : :
1935 CTCGGCTTCGCCATCACCCTCTCCACACACTGCTCCGGAGTGGGTC 1886
17 gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
||||| : : : : : ||||| : : : : :
1885 CCCCTTCAGCCAGGC.....CCACGCTCCCA 1857
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
||
1856 CT..... 1855
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
||| ||| ||| : : : ||||
1854 .....CCCCAGGACCCCTTGGCGCCCAAGCCACACCACC 1820
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
||||| : : : : : |||||
1819 TGTCTCCAGCCCATCTGGTCTCCACTGTGGCCCTG..... 1783
84 rgArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
: : : ||| ||| |||
1782 ..GCTCCGGCTGGCTCCACGGCACACCGCGGC..... 1750
101 Thr AspArgSerProAlaProMetSerCysAspLysSerThrGln...T 116
||| : : : |||| ||| ||| ||| : : : : : |||
```


OM of: US-09-508-832-10 to: Issued_Patents_NA:* out_format : pfs

Date: Aug 7, 2002 10:36 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framer_p2n.spool -DEV=xlh
-O/cgn2_1/USPTO.spool/US09508832/runat_06082002_103611_2799/app_query.fasta_1.259
-DB=Issued_Patents_NA -QWTF=fastap -SUFTX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -QCAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -QCAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=humana40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFT=pfs
-NORM=ext -HEAPSIZE=500 -NINLEN=0 -MAXLEN=2000000000
-USBR=US09508832.ecgn1_1.46 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-508-832-10
Query length: 198
Database: Issued_Patents_NA:*
Database sequences: 381533
Database length: 122816752
Search time (sec): 46.200000

score list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-474-379C-64	111.50	147.06	1.67	3705	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-146-249A-64	111.50	147.06	1.67	3705	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-206-188B-64	111.50	147.06	1.67	3705	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-080-115	107.00	158.63	0.3796	647	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-362-115	107.00	158.63	0.3796	647	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-095-855-115	107.00	158.63	0.3796	647	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-324-542-115	107.00	158.63	0.3796	647	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-602-359A-24	107.00	153.54	0.7291	1041	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-080-113	107.00	149.15	1.28	1569	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-362-113	107.00	149.15	1.28	1569	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-095-855-113	107.00	149.15	1.28	1569	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-324-542-113	107.00	149.15	1.28	1569	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-080-159	107.00	148.77	1.34	1626	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-362-159	107.00	148.77	1.34	1626	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-095-855-159	107.00	148.77	1.34	1626	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-324-542-159	107.00	148.77	1.34	1626	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-144-085-3	106.50	115.60	94.63	33529	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-105-537-9	106.00	148.35	1.42	1458	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-320-878-24	106.00	147.60	1.56	1565	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-135-994-2	106.00	145.73	1.99	1864	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-105-537-3	106.00	124.45	30.40	13613	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-998-416-1077	104.50	156.02	0.5306	571	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-484-840-1	104.00	143.09	2.79	1776	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-483-094-1	104.00	143.09	2.79	1776	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-09-318-448-13	104.00	143.09	2.79	1776	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-103-840A-2	104.00	59.46	1.2e+05	4403765	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-09-103-840A-1	104.00	59.44	1.2e+05	4411529	
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-764-233A-1	103.00	105.93	326.99	49377	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-335-409-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-568-102-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-567-969-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-568-480-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-568-486-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-568-472-1	102.50	101.60	569.85	68750	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-443-501A-2	102.00	100.32	671.70	71989	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-486-273A-44	101.50	132.70	10.56	3243	
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-486-273A-44	101.50	132.70	10.56	3243	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-480-474-44	101.50	132.70	10.56	3243	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-940-086A-44	101.50	132.70	10.56	3243	
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-940-035A-44	101.50	132.70	10.56	3243	

/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-231-193A-43 + 101.50 131.29 12.65 3698
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-486-273A-43 + 101.50 131.29 12.65 3698
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-480-474-43 + 101.50 131.29 12.65 3698
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-940-086A-43 + 101.50 131.29 12.65 3698
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-940-035A-43 + 101.50 131.29 12.65 3698

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-474-379C-64

seq_documentation_block:
: Sequence 64, Application US/08474379C
: Patent No. 5977305
: GENERAL INFORMATION:
: APPLICANT: Wigler, Michael H.
: APPLICANT: Colicelli, John J.
: TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
: TITLE OF INVENTION: PROCESSES
: NUMBER OF SEQUENCES: 88
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive/6300 Sears Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474,379C
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US-07/511,715
: FILING DATE: 20-APR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/206,188
: FILING DATE: 01-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/688,352
: FILING DATE: 19-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Clough, David W.
: REGISTRATION NUMBER: 36,107
: REFERENCE/DOCKET NUMBER: 27866/32771
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: INFORMATION FOR SEQ ID NO: 64:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3705 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE: CDS
: NAME/KEY: CDS
: LOCATION: 116..2773
: US-08-474-379C-64

alignment_scores:
Quality: 111.50 Length: 171
Ratio: 1.282 Gaps: 7
Percent Similarity: 50.877 Percent Identity: 27.485

alignment_block:

US-09-508-832-10 x US-08-474-379C-64
Align seg 1/1 to: US-08-474-379C-64 from: 1 to: 3705
6 SeraspvalserSerGlcysaspargGluGlyargGlnLeuGlnProAl 22

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic RNA
 US-08-997-080-115

alignment_scores:

Quality: 107.00 Length: 162
 Ratio: 1.259 Gaps: 9
 Percent Similarity: 52.469 Percent Identity: 30.864

alignment_block:

US-09-508-832-10 x US-08-997-080-115

Align seg 1/1 to: US-08-997-080-115 from: 1 to: 647

```

17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAATTGGGTATGACGAAGAGCGCCGCTGGCTCGAGCGGGGCTC 57
|||||
33 OthSerLeuGlnThrGluProGlnGlyAsnPro GluGlyAsnHisGly 49
|||||
58 AACGCCCTCGCAGAGC .....CCGTAAAGGTGACGTTGGG 92
|||||
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAGCGCTCGCAACGTCGCTGCGAAGAAAGTGGGGCG .....CC 135
|||||
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCAGCATCACCA...ACGATGCTGTGTCATCGCCA..... 169
|||||
83 etArgArgSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ...AGGATCGAGCTGGAGGACCGCTACGAGAGATCGCGCTGAGCTGG 217
|||||
100 AspThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnTh 116
|||||
218 ...TCAAAGAGTGGCGAGAGACGACGACGTCGCGGGGCGGCGCACC 264
|||||
116 ProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||
265 ACCACCGCCACCGTCGCTCGCTCAGGCTCTGCTCGCAAGGCTCGCAA 314
|||||
133 erMetArgGlnAlaGluProAlaAspMetArgProGluIleTrrp..... 147
|||||
315 CGTCCGACCGCGCCCAACCGCT .....CGGCTCAAGCTGCGCATCG 358
|||||
148 .....IleAlaGlnGluLeuArgTleg 156
|||||
359 AGAAGGCTGTGAGGCTGTCAACCGATCGCTGCTCAAGTCGCGCAAGGAG 408
|||||
156 lyAspGluPheAsnAlaTyrTalaArg 166
|||||
409 GTCGAGACCAAGGACGAGATTCTGCCACCG 440

```

seq_name: /cgn2_6/ptodata/1/lna/58_COMB.seq:US-08-997-362-115

seq_documentation_block:

Sequence 115, Application US/08997362

Patent No. 5985287

GENERAL INFORMATION:

APPLICANT: Tan, Paul
 APPLICANT: Hiyama, Jun
 APPLICANT: Visser, Elizabeth
 APPLICANT: Skinner, Margot
 APPLICANT: Scott, Linda
 APPLICANT: Prestidge, Ross
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
 TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
 NUMBER OF SEQUENCES: 194
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/997,362
 FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
 FILING DATE: June 12, 1997
 APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
 FILING DATE: August 29, 1996
 ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet
 REGISTRATION NUMBER: 37,007
 REFERENCE/DOCKET NUMBER: 11000.1002c2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-269-0565
 TELEFAX: 206-269-0563
 TELEX:

INFORMATION FOR SEQ ID NO: 115:

SEQUENCE CHARACTERISTICS:
 LENGTH: 647 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic RNA
 US-08-997-362-115

alignment_scores:

Quality: 107.00 Length: 162
 Ratio: 1.259 Gaps: 9
 Percent Similarity: 52.469 Percent Identity: 30.864

alignment_block:

US-09-508-832-10 x US-08-997-362-115

Align seg 1/1 to: US-08-997-362-115 from: 1 to: 647

```

17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAATTGGGTATGACGAAGAGCGCCGCTGGCTCGAGCGGGGCTC 57
|||||
33 OthSerLeuGlnThrGluProGlnGlyAsnPro GluGlyAsnHisGly 49
|||||
58 AACGCCCTCGCAGAGC .....CCGTAAAGGTGACGTTGGG 92
|||||
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAGCGCTCGCAACGTCGCTGCGAAGAAAGTGGGGCG .....CC 135
|||||
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCAGCATCACCA...ACGATGCTGTGTCATCGCCA..... 169
|||||
83 etArgArgSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ...AGGATCGAGCTGGAGGACCGCTACGAGAGATCGCGCTGAGCTGG 217
|||||
100 AspThrAspArgSerProAlaPrometSerCysAspLysSerThrGlnTh 116
|||||
218 ...TCAAAGAGTGGCGAGAGACGACGACGTCGCGGGGCGGCGCACC 264
|||||
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||

```


Percent Similarity: 52.469 Percent Identity: 30.864

alignment_block:

US-09-508-832-10 x US-09-324-542-115 ..

Align seg 1/1 to: US-09-324-542-115 from: 1 to: 647

```

17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 AGACAAATGGGTATGACGAAGAGCGCCGCGTGGCCTCGAGCGGGCGTC 57
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 AACGCCCTCCAGACG.....CCGTAAAGGTGACGTTGGG 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93 CCCGAGGGTGGACACGCTGCTGCGAGAGAACTGGGCG.....CC 135
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
136 CCCAGATACCA...ACGATGCTGTCTCATCGCCA.....169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 eAlaArgSerSerLeuLeuSerAlaArgSerSerGlyTyPheSerPhe 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 ..ACGACATCGAGCTGGAGGACCGCTACGAGNAGATCGCGCTGACCTGG 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 116
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 ...TCGAAGAGGTGGCCAGACGACGACGCTGCGGGGCGGACGCC 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 rProSerProCysGlnAlaPheAsnHisTyLeuSerAlaMetAla 133
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 ACCACCGCCACCGTCTGCTCAGGCTCTGCTCGAAGGCGCTGCCAA 314
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp..... 147
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 CGTCGACGCGGCGCCAAACCGCT.....CGGCTCAAGCGTGGCATCG 358
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 .....IleAlaGlnGluLeuArgArgIleG 156
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
359 AGAAGGCTGTCGAGGCTGTACCCAGTCTGCTGAAGTGGGCCAAGGAG 408
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 lyAspGluPheAsnAlaTyTyAlaArgArg 166
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
409 GTCGAGACCAAGGACGAGATTTCTGCCACCGC 440

```

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-602-359A-24

seq_documentation_block:

Sequence 24, Application US/08602359A

Patent No. 5942430

GENERAL INFORMATION:

```

APPLICANT: ROBERTSON, Daniel E.
APPLICANT: MURPHY, Dennis
APPLICANT: REID, John
APPLICANT: MAFFIA, Anthony
APPLICANT: LINK, Steven
APPLICANT: SWANSON, Ronald V.
APPLICANT: WARREN, Patrick V.
APPLICANT: KOSMOTKA, Anna
TITLE OF INVENTION: ESTERASES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & RICHARDSON P.C.
STREET: 4225 EXECUTIVE SQUARE, STE 1400
CITY: LA JOLLA
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 92037

```

COMPUTER READABLE FORM:

```

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2

```

```

OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602.359A
FILING DATE: February 16, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: HAILE, LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1041 NUCLEOTIDES
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: GENOMIC DNA
US-08-602-359A-24

alignment_scores:
Quality: 107.00 Length: 177
Ratio: 1.244 Caps: 9
Percent Similarity: 48.588 Percent Identity: 27.684

```

alignment_block:

US-09-508-832-10 x US-08-602-359A-24/rev ..

Align seg 1/1 to reverse of: US-08-602-359A-24 from: 1 to: 1041

```

2 AlalysGlnProSerAspValSerSerGluCysAsp.....13
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
757 GCACACCCACCCACCGCGGCGGTCTTGTACTCCAGCCACCGGGGT 708
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14 ....ArgGluGlyArgGlnLeuGlnProAlaGluArgProGlnLeu 29
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
707 ATGACGTCCTCCGAGCCCTATAGTACGGGCTATCAGCAGCCGCGT...A 661
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 rGProGlyAla.....ProfhrSerLeuGlnThrGluProGln 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
660 GACCGCGGGTCCCGACCCCTCCACGATCGCTACAGCGCGGCCATAC 611
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
42 GlyAsnProGluGlyAsnHisGlyGlyAspSerCysProHisG 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
610 TGAACCCCAATATTATCCGGCGCGGGGAAACCGCTCCGAGACATAG 561
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 ySerProGlnGlyProLeuAlaProAlaSerProGlyProPheAla 75
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
560 CCCACACAGGCC.....GGCATCCAGCACCTCCCGGGGCCAATCGT 517
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
75 hrArgSerProLeuPheIlePheMetArgSerSerLeuLeuSerArg 91
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516 COTCAGACCCC.....CGC 503
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
92 SerSerSerGlyTyPheSer.....PheAspThrAs 102
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
502 TCTCCCGTGGCCCGGAAGTCGAACACACCGGGTACCCCGCATCC 453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
102 pArgSerProAlaProMetSerCysAspLysSerThrGln.ThrProSer 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
452 ACGAGCTCCCGGGCCAGCACGCGCATGTAGGGCGCGGCGGCCACCCAGT 403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 ProProCysGlnAlaPheAsnHisTyLeuSerAlaMetAlaSerMetAr 135
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
402 ATACCCGTGCATCAAAACGAACACCGGGTTCGCCCGCGCGGGGCCCA 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
135 gclnAlaGluPro.....AlaAspMetArg.....ProGluIleT 147

```



```
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1569 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-997-362-113

alignment_scores:
  Quality: 107.00      Length: 162
  Ratio: 1.259        Gaps: 9
  Percent Similarity: 52.469  Percent Identity: 30.864

alignment_block:
US-09-508-832-10 x US-08-997-362-113 ..
Align seg 1/1 to: US-08-997-362-113 from: 1 to: 1569
17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAATTGGCTATGACGAAGAGCCCGCTGCGCTCGAGCGGGCCTC 57
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||||
58 AACGCCCTCGCAGACG.....CCGTAAAGGTGACGTTGGG 92
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAAGGCTCGCAACGCTGCTGCTGGAAGAAGTGGCGG.....CC 135
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCACGATCACCA...ACGATGGTGTCTCATCCGCA..... 169
83 eArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ..AGGAGTACGAGCTCGAGGACCCGTCACGAGAAGATCGGCGCTGAGCTGG 217
100 AspThrAspArgSerProAlaProMetSerCysAspHisSerThrGlnTh 116
...
218 ...TCMAAGAGTTCGCGACAGACACGACGCTCGCGGCGGCGACAC 264
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||
265 ACCACCGCCACGCTGCTGCTGAGGCTGCTGCGGAGGCGCTGCGCAA 314
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp..... 147
315 GTCGACAGCGCGCCCAACCCGCT.....CGGCTCAAGCGTGGCATCG 358
148 .....
359 AGAAGGCTGTGAGGCTGTACCCAGTCGCTGCTGAGTCGCGCAAGGAG 408
156 lYAspGluPheAsnAlaTyrTyrAlaArgArg 166
409 GTCGAGACCAAGGAGGAGATTTCTGCCACCGC 440

seq_name: /cgn2_6/ptodata/1/1na/68_COMB.seq:US-09-095-855-113

seq_documentation_block:
; Sequence 113, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
```

```
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann M. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PestSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1569 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-09-095-855-113

alignment_scores:
  Quality: 107.00      Length: 162
  Ratio: 1.259        Gaps: 9
  Percent Similarity: 52.469  Percent Identity: 30.864

alignment_block:
US-09-508-832-10 x US-09-095-855-113 ..
Align seg 1/1 to: US-09-095-855-113 from: 1 to: 1569
17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGACAATTGGCTATGACGAAGAGCCCGCTGCGCTCGAGCGGGCCTC 57
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||||
58 AACGCCCTCGCAGACG.....CCGTAAAGGTGACGTTGGG 92
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAAGGCTCGCAACGCTGCTGCTGGAAGAAGTGGCGG.....CC 135
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCACGATCACCA...ACGATGGTGTCTCATCCGCA..... 169
83 eArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ..AGGAGTACGAGCTCGAGGACCCGTCACGAGAAGATCGGCGCTGAGCTGG 217
```



```
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 ...TCAAAGAGTCCGCAAGAACACGACGTCGCGGCGGACGGCACC 264
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
265 ACCACCGCCACCGTCTGCTCAGGCTCTGCTCGCGAAGGCGTCGCA 314
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp.... 147
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
315 CGTCGACGCGCGCCCAACCGCT.....CGGCCTCAAGCGTGGCATCG 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 .....IleAlaGlnGluLeuArgArgileG 156
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
359 AGAAGCGTCTCGAGCGTGTCCACCGACGCTGCTGAAGTCGCCAAGGAG 408
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 lyAspGluPheAsnAlaTyrTyrAlaArg 166
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409 GTCGAGACCAAGGACGAGATTCTGCCACCGC 440
```

seq_name: /cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-324-542-113

```
seq_documentation_block:
; Sequence 113, Application US/09324542
; Patent No. 5328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
;   OF IMMUNOLOGICALLY-MEDIATED SKIN DISORDERS
; FILE REFERENCE: 11000.1007c1
; CURRENT APPLICATION NUMBER: US/09/324.542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 113
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-324-542-113
```

alignment_scores:
Quality: 107.00 Length: 162
Ratio: 1.259 Gaps: 9
Percent Similarity: 52.469 Percent Identity: 30.864

alignment_block:

US-09-508-832-10 x US-09-324-542-113 ..

Align seq 1/1 to: US-09-324-542-113 from: 1 to: 1569

```
17 ArgGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaPr 33
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
8 ACACAAATTCCTATGACGAGAGCCGCGCGTGGCTCGACGGCGGCGCTC 57
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
33 othrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
58 AACGCCCTCGGACG.....CCGTAAAGGTGACGTTGGG 92
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
93 CCGGAAGGGTCGCAACGTCGCTGGAGAGAAGTGGGGCG.....CC 135
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
136 CCCACGATCACCA...ACGATGGTGTTCATCGCCA..... 169
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
83 eArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
170 ..AGGACATCGAGCTGGAGGACCGTACGAGAGATCGGCGCTGAGCTGG 217
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 ...TCAAAGAGTCCGCAAGAACACGACGTCGCGGCGGACGGCACC 264
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
265 ACCACCGCCACCGTCTGCTCAGGCTCTGCTCGCGAAGGCGTCGCA 314
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTrp..... 147
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
315 CGTCGACGCGCGCCCAACCGCT.....CGGCCTCAAGCGTGGCATCG 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 .....IleAlaGlnGluLeuArgArgileG 156
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
359 AGAAGCGTCTCGAGCGTGTCCACCGACGCTGCTGAAGTCGCCAAGGAG 408
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 lyAspGluPheAsnAlaTyrTyrAlaArg 166
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409 GTCGAGACCAAGGACGAGATTCTGCCACCGC 440
```

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-997-080-159

seq_documentation_block:

```
; Sequence 159, Application US/08997080
; Patent No. 5968524
```

```
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
;   MEDIATED SKIN DISORDERS
; FILE REFERENCE: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 159:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

US-08-997-080-159

alignment_scores:
Quality: 107.00 Length: 162
Ratio: 1.259 Gaps: 9
Percent Similarity: 52.469 Percent Identity: 30.864

alignment_block:

US-09-508-832-10 x US-08-997-080-159

Align seg 1/1 to: US-08-997-080-159 from: 1 to: 1625

```

17 ArgInLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGAAATTCGTATGACGAAGAGCCGCCCTGCGCTCGAGCGGGCCTC 57
|||||
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||
58 AACGCCCTCGCAGACG.....CCGTAAGGTGACGTGGG 92
|||||
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAAGGTCGACGCTGCTCGAGAGAACGTGGGCG.....CC 135
|||||
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCACGATCACCA...ACGATGCTGTCTCCATCCCA..... 169
|||||
83 eTArgArgSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ..AGGAGATCGACCTGGAGGACCCGTACGAGAAGATCGCGCTGAGCTG 217
|||||
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 116
|||||
218 ...TCAAAGAGTCCGACGAGACCGACGCTCGCGGCGAGCGCAC 264
|||||
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||
265 ACCACGCCACCGCTGCTCGAGCTCTGCTCGGAGGCTCGGCA 314
|||||
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTyr..... 147
|||||
315 CGTCGACGCGCGCAACCCGCT.....CGGCTCAAGCTGGCATCG 358
|||||
148 .....IleAlaGlnGluLeuArgArgIleG 156
|||||
359 AGAGGCTGTCGAGGCTGTACCCAGTCTGCTGAAGTCGGCAAGGAG 408
|||||
156 lyAspGluPheAsnAlaTyrTyrAlaArgArg 166
|||
409 GTCGAGACCAAGGACGAGATTCTGCCACCGC 440

```

seq_name: /cgn2_6/ptodata/1/ins/5B_COMB.seq:us-08-997-362-159

seq_documentation_block:

Sequence 159 Application US/08997362

Patent No. 5985287

GENERAL INFORMATION:

```

APPLICANT: Tan, Paul
APPLICANT: Hiyaama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross

```

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR

TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS

NUMBER OF SEQUENCES: 194

CORRESPONDENCE ADDRESS:

ADDRESSES: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/997.362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002C2
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 1626 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-362-159

```

alignment_scores:

Quality: 107.00 Length: 162

Ratio: 1.259 Gaps: 9

Percent Similarity: 52.469 Percent Identity: 30.864

alignment_block:

US-09-508-832-10 x US-08-997-362-159

Align seg 1/1 to: US-08-997-362-159 from: 1 to: 1625

```

17 ArgGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPr 33
|||||
8 AGAAATTCGTATGACGAAGAGCCGCCCTGCGCTCGAGCGGGCCTC 57
|||||
33 oThrSerLeuGlnThrGluProGlnGlyAsnPro.GluGlyAsnHisGly 49
|||
58 AACGCCCTCGCAGACG.....CCGTAAGGTGACGTGGG 92
|||||
50 GlyGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaPr 66
|||||
93 CCCGAAGGTCGCAACGCTGCTCGAGAGAACGTGGGCG.....CC 135
|||||
66 oProAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheM 83
|||||
136 CCCACGATCACCA...ACGATGCTGTCTCCATCCCA..... 169
|||||
83 eTArgArgSerLeuLeuSerArgSerSerGlyTyrPheSerPhe 99
|||||
170 ..AGGAGATCGACCTGGAGGACCCGTACGAGAAGATCGCGCTGAGCTG 217
|||||
100 AspThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnTh 116
|||||
218 ...TCAAAGAGTCCGACGAGACCGACGCTCGCGGCGAGCGCAC 264
|||||
116 rProSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlas 133
|||||
265 ACCACGCCACCGCTGCTCGAGCTCTGCTCGGAGGCTCGGCA 314
|||||
133 erMetArgGlnAla.GluProAlaAspMetArgProGluIleTyr..... 147
|||||
315 CGTCGACGCGCGCAACCCGCT.....CGGCTCAAGCTGGCATCG 358
|||||
148 .....IleAlaGlnGluLeuArgArgIleG 156
|||||
359 AGAGGCTGTCGAGGCTGTACCCAGTCTGCTGAAGTCGGCAAGGAG 408
|||||
156 lyAspGluPheAsnAlaTyrTyrAlaArgArg 166
|||

```


ORIGIN

alignment_scores: Quality: 667.00 Length: 146
 Ratio: 4.869 Gaps: 2
 Percent Similarity: 93.836 Percent Identity: 89.726

alignment_block:
 US-09-508-832-10 x BB651225 ..

Align seg 1/1 to: BB651225 from: 1 to: 645

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGly 17
 |||||
 220 ATGCCAAGCAGCACTTCGTGATGTAAGTCTGAGTGTGACAGAGAAGGTGG 269
 |||||
 17 -gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34
 |||||
 270 ACAATTGTCAGCCTGCTGACAGAGCCCTCCCGAGCTCAGGCGCTGGGCCCTTA 319
 |||||
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
 |||||
 320 CCTCCCTACAGACAGAACCGCAGGTATCCGAC.....GGC 357
 |||||
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
 |||||
 358 GAAGGGAGCGCTGGCCCCACGGCAGCCCTCAGGGGCCCGCTGGCCCCCACC 407
 |||||
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
 |||||
 408 GGCCAGCCCTGGCCCTTTTGTCTACAGATCCCGACATTTTCATCTTTGTGA 457
 |||||
 84 rGArgSerSerLeuSerArgSerSerGlyTyPheSerPheAsp 100
 |||||
 458 GAAGHCTCTCTGCTGCTCCCGCTCTCCAGTGGTAATCTCTTTTGAC 507
 |||||
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
 |||||
 508 ACAGACAGGAGCGCGCCACCCATGATGTTGTGACAAAGTCAACACAAACCCC 557
 |||||
 117 oSerProProCysGlnAlaPheAsnHisTyPheLeuSerAlaMetAlaSerM 134
 |||||
 558 AAGTCTCTCTGCGAGCGCTTCCAC.CACTATCTCAGTCAATGGCTTCCA 606
 |||||
 134 eTArgGlnAla.....GluProAlaAspMetArgPro 144
 |||||
 607 TACGACAGTCTCAGGAGGAACCTGAAGATCTGCCGCCG 644

seq_name: gb_gss:A2706148

seq_documentation_block:

LOCUS A2706148 580 bp DNA linear GSS 24-JAN-2001

DEFINITION RPCI-23-227P3.TV RPCI-23 Mus musculus genomic clone RPCI-23-227P3,

ACCESSION DNA sequence.

VERSION A2706148

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 580)

AUTHORS Zhao,S., Niezman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret

,B., Levins,M., McGann,S., Tsagaye,G., Geer,K., Krol,M., de Jong,P.

and Fraser,C.M.

TITLE Mouse BAC End Sequences from Library RPCI-23

JOURNAL unpublished (1999)

COMMENT Other_GSSs: RPCI-23-227P3.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200
 Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tcldb/bac_ends/mouse/bac_ends_intro.html

Plate: 227 row: P column: 3

Seq primer: 17

Class: BAC ends.

FEATURES

source

1..580 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-227P3"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site: 1:
 EcoRI; Site: 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."
 BASE COUNT 138 a 162 c 138 g 142 t
 ORIGIN

alignment_scores:

Quality: 656.00 Length: 131

Ratio: 5.248 Gaps: 1

Percent Similarity: 95.420 Percent Identity: 93.893

alignment_block:

US-09-508-832-10 x A2706148 ..

Align seg 1/1 to: A2706148 from: 1 to: 580

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGly 17

90 ATGCCAAGCAACCTTCGTGATGTAAGTCTCTGAGTCTGACAGAGAAGGTGG 139

17 -gGlnLeuGlnProAlaGluArgProProGlnLeuArgProGlyAlaProT 34

140 ACAATTGTCAGCCTGCTGACAGAGCCCTCCCGAGCTCAGGCGCTGGGCCCTTA 189

34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50

190 CCTCCCTACAGACAGAACCGCAGGTATCCCGAC.....GGC 227

51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67

228 GAAGGGAGCGCTGCCCGCAGCGCCCTCAGGCCCGCTGGGCCCCACC 277

67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84

278 GGCCAGCCCTGGCCCTTTTGTCTACAGATCCCGACATTTTCATCTTTGTGA 327

84 rGArgSerSerLeuSerArgSerSerGlyTyPheSerPheAsp 100

328 GAAGTCTCTCTGCTGCTCCCGCTCTCCAGTGGGTATTTCTCTTTTGAC 377

101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117

378 ACAGACAGGAGCGCGCCACCCATGATGTTGTGACAAAGTCAACACAAACCCC 427

117 oSerProProCysGlnAlaPheAsnHisTyPheLeuSerAlaMet 131

428 AAGTCTCTCTTGGCAGGCGCTTCAACCACTATCTCAGTCAATG 470

```

seq_name: gb_est1:BB577135
seq_documentation_block:
LOCUS      BB577135          586 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION BB577135 RIKEN full-length enriched, 13 days embryo male testis Mus
            musculus cDNA clone 6030401M20 5', mRNA sequence.
ACCESSION  BB577135
VERSION    BB577135.2 GI:16449433
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 686)
AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
            Hiramoto,K., Hori,Y., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakai,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
            RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
            Unpublished (2001)
JOURNAL   On Nov 30, 2000 this sequence version replaced gi:11473679.
COMMENT   Contact: Yoshinori Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp
            URL: http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh
            ,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
            Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Location/Qualifiers
            1. 686
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone_lib="RIKEN full-length enriched, 13 days embryo
            male testis"
            /sex="male"
            /tissue_type="testis"
            /dev_stage="13 days embryo"

FEATURES
            source
            seq_name: gb_est2:BI906766
            seq_documentation_block:
            LOCUS      BI906766          700 bp      mRNA      linear      EST 16-OCT-2001
            DEFINITION 603064506F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5213713 5',
            mRNA sequence.
            ACCESSION  BI906766
            VERSION    BI906766.1 GI:16169524
            KEYWORDS   EST.

```

```

/lab_host="DH10B"
/note="Site_1: Sall; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GACGAGACGACGACGACGACGCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 5.0 and subtraction to Rot = 100.0. Second strand
cDNA was prepared with the primer adapter of sequence [5'
GACGAGACGACGACGACGACGCTTTTCTTTTNN 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI."
BASE COUNT      145 a      228 c      167 g      146 t
ORIGIN
alignment_scores:
            Quality: 653.00      Length: 131
            Ratio: 5.224      Gaps: 1
            Percent Similarity: 95.420      Percent Identity: 93.130
alignment_block:
US-09-508-832-10 x BB577135 ..
Align seg 1/1 to: BB577135 from: 1 to: 686
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
220 ATGCCCAACCAACCTCTGAGTGAAGTCTGAGTGCACAGAGAAGGTGG 269
17 gGlnLeuGlnProAlaGluArGProProGlnLeuArGProGlyAlaProT 34
270 ACAATTCCAGCCCTGCTGAGAGGGCCCTCCAGCTCAGGCGCTGGGGCCCTA 319
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
320 CCTCCCTACAGACAGACCCGCAAGGTATCCGAC.....GGC 357
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
358 GAAGGGGACCGCTGGCCCGCAGCGCCCTCAGGGCCCGCTGGCCCCACC 407
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
408 GGCCAGCCCTGGCCCTTTTGTCTACAGATCCCACTTTTCATCTTTGTGA 457
84 rArgSerSerLeuLeuSerArgSerSerSerGlyTyrPheSerPheAsp 100
458 GAAGATCTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 507
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
508 ACAGACAGAGGCGCGCCACCCCTGAATTTGTGACAAGTCAACACAAACCC 557
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMet 131
558 AAGTCCTCTCTTGGCCGCTTCAACCACTATCTCATGTGCAATG 600
seq_name: gb_est2:BI906766
seq_documentation_block:
LOCUS      BI906766          700 bp      mRNA      linear      EST 16-OCT-2001
DEFINITION 603064506F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5213713 5',
            mRNA sequence.
            ACCESSION  BI906766
            VERSION    BI906766.1 GI:16169524
            KEYWORDS   EST.

```


SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 700)
TITLE NIH-MGC http://mgs.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1536 row: 1 column: 02
High quality sequence stop: 696.
seq_name: gb_est2:BF021882
seq_documentation_block:
LOCUS BF021882
DEFINITION uy59b09.v1 McCartney Eddy round spermatid Mus musculus cDNA clone
IMAGE:3663833 5' similar to TR:054918 054918 BCL2 INTERACTING
MEDIAN OF CELL DEATH ;, mRNA sequence.
ACCESSION BF021882
VERSION BF021882.1 GI:10753214
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 452)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, I., Person,
B., Svaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter,
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R., and Wilson, R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1424601
Seq primer: Primer name ambiguous
High quality sequence stop: 386.
FEATURES
Location/Qualifiers
1..452
/organism="Mus musculus"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:3663833"
/clone_lib="McCarrey Eddy round spermatid"
/sex="male"
/tissue_type="round spermatids, pooled from multiple mice"
/dev_stage="60 day"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pBluescript SK+ (Stratagene
); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dt-primed
[5'-(CA)10-ACTAGTCTCGAGTTTITTTT-3'] and directionally
cloned using 5' linkers 5'-AATTGGCAGGAG-3' and
5'-CTCGTCCG-3'. Size selection of >400bp material gives
average insert size ranging from 1-2 kb. Library was mass
excised (from lambda-unizap-XR) and resulting
single-stranded phagemids were prepped and transformed
into DH10B. Library contains 98.5% recombinants.
References: J. Androl. 20:635-639 and Gene 25:263-269.
Library constructed and donated by J. McCarrey, Ph.D.
(Southwest Foundation for Biomedical Research, Dept. of

BASE COUNT 148 a 233 c 175 g 144 t
ORIGIN
alignment_scores:
Quality: 651.50 Length: 151
Ratio: 4.402 Gaps: 2
Percent Similarity: 98.013 Percent Identity: 95.364
alignment_block:
US-09-508-832-10 x B1906766 ..
Align seg 1/1 to: B1906766 from: 1 to: 700
1 MetAlaLysGlnProSerAspValSerSer.GluCysAspArgGluGlyA 17
257 ATGGCAAGCAACCTTCTGATGTAAGTTCTGAGTGTGACCGAGAGGTA 306
17 rGgInLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaPro 33
307 GACAAATTGCAGCGTCGGAGAGCGCTCCAGCTCAGACCTGGGGCCCT 356
34 ThrSerLeuGlnThrGlnProGlnGlnArgProGluGlyAsnHisGlyGI 50
357 ACCCTCCCTACAGACAGAGCCACAGGATATCTGAAGCAATCAGGAGG 406
50 yCluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProp 67
407 TCAAGGGGACAGCTGCCCGCCAGGAGCC...TCAGGGCCCGTGGCCCGAC 453
67 roAlaSerProGlyProPheAlaThrArgSerProLeuGlyPheIlePheMet 83
454 CTGGCCAGCCCTGG. CTTTGTGTACAGATCCCGCTTTTCATCTTATG 502
84 ArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAs 100
503 AGAAGATCCTCCCTGGTGTCTGATCCTCCAGTGGGTATTC.TCTTTTGA 551

Genetics): excision done by E.M. Eddy, Ph.D. (National Institutes of Health, National Institute of Environmental Health Sciences). Original lambda-based library is available through ATCC, catalog #63423.

BASE COUNT 106 a 130 c 112 g 104 t
ORIGIN

alignment_scores:
Quality: 622.00 Length: 139
Ratio: 4.822 Gaps: 1
Percent Similarity: 92.806 Percent Identity: 85.612

alignment_block:
US-09-508-832-10 x BF021882 ..
Align seg 1/1 to: BF021882 from: 1 to: 452

```

62 GlyProLeuAlaProProAlaSerProGlyProPheAlaThrArgSerPr 78
1 GGGCGGCTGGCCGCCACGGCGACGCTTGGCCCTTTGCTAGCAGATCCCC 50
78 oLeuPheMetArgSerSerLeuLeuSerArgSerSerG 95
51 ACTTTTCATCTTGTGAGAGATCTTCTGCTGCTGCCGGTCTCCAGTG 100
95 lYTyPheSerPheAspThrAspArgSerProAlaProMetSerCysAsp 111
101 GGTATTTCTCTTTTGACACAGACGAGCGCGGCACCATGAGTGTGAC 150
112 LysSerThrGlnThrProSerProProCysGlnAlaPheAsnHisTyrLe 128
151 AAGTCAACACAAACCCCAAGTCTCTCTGCCAGGCTTCAACCACTATCT 200
128 uSerAlaMetAlaSerMetArgGlnala.....GluProAlaAspMetA 143
201 CAGTGCAATGCTTCCATACGACGCTCTCAGGAGGAACCTGAAGATCTG 250
143 rProGluLeuTrpPheAlaGlnGluLeuArgArgIleGlyAspGluPhe 159
251 GCGCGGAGATACGGATTGACACAGAGAGTGGCGGATCGGAGAGGATTC 300
160 AsnAlaTyrTrpAlaArgArgValPheLeuAsnAsnTyrGlnAlaAlaG 176
301 AACGAACCTTACACAGAGGCTGTTTGCANATGATTACCGCGAGGCTGA 350
176 uAspHisProArgMetValIleLeuArgLeuLeuArgTyrIleValArgL 193
351 AGACCAACCTCAATGTTATCTTACACTGTACGCTTTATCTTCCTGC 400
193 euValTrpArgMethis 198
401 TGGTATGGAAGGCAT 417

```

seq_name: gb_est2:BI454637

seq_documentation_block:
LOCUS BI454637 765 bp mRNA linear EST 21-AUG-2001
DEFINITION 603170532F1 NCI_CCAP_Mam5 Mus musculus cDNA clone IMAGE:5250271 5',
mRNA sequence.

ACCESSION BI454637
VERSION BI454637.1 GI:15245293
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE
1 (bases 1 to 765)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999);
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
cDNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L14M11631 row: 0 column: 08
High quality sequence stop: 760.

FEATURES
source

Location/Qualifiers
1..765
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone_lib="NCI_CCAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 175 a 214 c 227 g 149 t
ORIGIN

alignment_scores:
Quality: 576.00 Length: 201
Ratio: 4.465 Gaps: 3
Percent Similarity: 64.179 Percent Identity: 59.701
alignment_block:
US-09-508-832-10 x BI454637 ..
Align seg 1/1 to: BI454637 from: 1 to: 765
1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
172 ATGCCCAAGCAACCTTCTGATGTAAGTTCGTGAGTGTGACAGAGAGTGG 221
17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
222 ACANTTGCAGCCTGCTGAGAGGCTTCCACCTCAGGCTGGGGCCCCCA 271
34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
272 CCTCCCTACAGACAGAACGCCAA..... 294
51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
294 294
67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMetA 84
294 294
84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheAsp 100
294 294
101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
295 ...GACAGAGCGCGCACCCATGATGTTGTGACAAAGTCACACAAACCCC 341
117 oSerProProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
342 AACTCTCTCTGCCAGGCTTCAACCACTATCTCAGTGCATCGCTTCCA 391
134 etArgGlnAla.....GluProAlaAspMetArgProGluIleTrpIle 148
392 TAGCAGACTCTCAGGAGGAACCTGAAGATCTCGCGCGGAGATACCGATT 441

149 AlaGlnLeuLeuArgAlaGly AspGluPheAsnAlaTyrTyrAla 165
 442 GCACAGAGCTGGGGATCGGACGACGAGTTCACAAACTTACACAA 491
 165 rGargvalPheLeuAsnAsnTyrGlnAlaAlaGluAspHisProArgMet 181
 492 GGAGGGTGTTCAAATGATTACGCGAGGCTGAAGACCACTCAATG 541
 182 ValIleLeuArgLeuLeuArgTyrIleValArgLeuValIlePargMet 198
 542 GTTATCTTACAACTGTACGCTTATCTTCCTGCTGATGAGAGGCA 591
 198 s 198
 592 T 592

seq_name: gb_est2:BG173095

seq_documentation_block:
 LOCUS BG173095 668 bp mRNA linear EST 06-FEB-2001
 DEFINITION 60233666F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4459720 5',
 mRNA sequence.
 ACCESSION BG173095
 VERSION BG173095.1 GI:12679707
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NIH-MGC <http://mgs.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: rcpsp-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 plate: LLAM10260 row: c column: 17
 High quality sequence stop: 599.
 Location/Qualifiers

FEATURES

source
 1..668
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4459720"
 /clone_lib="NCI_CGAP_Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 135 a 235 c 162 g 136 t
 ORIGIN

alignment_scores:
 Quality: 544.50 Length: 154
 Ratio: 4.125 Gaps: 8
 Percent Similarity: 85.714 Percent Identity: 77.273

alignment_block:

US-09-508-832-10 x BG173095

Align seg 1/1 to: BG173095 from: 1 to: 668

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
 223 ATGCCAAGCAACCTTCTGATGTAGTCTGAGTGTGACAGAGAGGTGG 272
 17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProT 34
 273 ACAATTCACGCCCTGCTCAGAGCCCTCCAGCTCAGGCTGGGCCCTTA 322
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
 323 CCTCCTACAGACAGAACCCCAAGGTAAATCCGACGGC 360
 51 GluGlyAsp_SerCysProHisGlySerProGlnGlyProLeuAlaPro 67
 361 GAAGGGACCTGTGCCCCCAGCCCTCAGGGCCCTGCTGGCCCCCA 410
 67 roAlaSerProGlyProPheAlaThrArgSer_ProLeuPheHellePheMe 83
 411 CCGGACGCCCTGGCCCTTTTGCTACCAGATCCCCACCTTTTCATCTTTGT 460
 83 tArgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSer...P 99
 461 CAGAAATATCTCTCTGCTCCGCTCTCCAGTGGGTATATCTCCTT 510
 99 heAspThrAspArgSerProAlaProMetSerCysAspLysSerThrGln 115
 511 TTGACACAGCACAGGAGCCCGCCACCCATGAGTGTGACAACTCAACCAA 560
 116 ThrProSerProProCysGlnAlaPheAsnHisTyrLeuSer_AlaMet 131
 561 AACCCCAAGTCTCTTGGCAGGCCTTCAACCACTATCTCAGTTGCAATG 610
 132 AlaSerMetArgGln.....Ala_GluProAlaAspMetArgProGluI 146
 611 GCTTTCATACAGAGTCTCCAGGAGGAACCTGAGGATCTGCGCCCGGAGA 660
 146 le 146
 661 TC 662

seq_name: gb_est1:BB653788

seq_documentation_block:

LOCUS BB653788 684 bp mRNA linear EST 26-OCT-2001
 DEFINITION BB653788 RIKEN full-length enriched, adult male liver tumor Mus
 musculus cDNA clone C730026B05 5', mRNA sequence.

ACCESSION BB653788

VERSION BB653788.1 GI:16487616

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 684)

AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,

Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okada

Okazaki,Y., Okido,T., Saico,R., Sakai,C., Sakai,K., Sano,H., Sasaki

,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,

Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

TITLE

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome.res@gsc.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh

Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Prepare full-length cDNA libraries for cap-trapper-selected cDNAs to
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.

FEATURES
 source
 1. 480
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G37003SM16"
 /cell_type="B16 F10Y cells"
 88 a 175 c 121 g 96 t

BASE COUNT
 ORIGIN

alignment_scores
 Quality: 376.00 Length: 88
 Ratio: 4.700 Gaps: 2
 Percent Similarity: 90.909 Percent Identity: 88.636

alignment_block:
 US-09-508-832-10 x BB856627 ..

Align seg 1/1 to: BB856627 from: 1 to: 480

1 MetAlaTysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
 |||||
 230 ATGCCAAGCAACCTTCTGATGTTCTAGTTCTGAGTGTACAGAGAGGTGG 279
 |||||
 17 gClnLeuGlnProAla.GluArgProProGlnLeuArgProGlyAlaPro 33
 |||||
 280 ACAATTGCAGCTGCTGTAGAGGCTCCCCAGCTCAGGCTGGGCCCT 329
 |||||
 34 ThrSerLeuClnThrGlnProGlnGlyAsnProGluGlyAsnHisGlyG 50
 |||||
 330 ACCTCCCTACAGACAGACG.CAAGGTAATCCGAC.....GG 366
 |||||
 50 yGluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProp 67
 |||||
 367 CGAAGGGGACCGCTGGCCCCAGGCGCCCTCAGGGCCGCTGGCCCCC 416
 |||||
 67 roAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMet 83
 |||||
 417 CGGCCAGCCCTGGCCCTTTTGTGTACAGATCCCCACCTTTTCATCTT 466
 |||||
 84 ArgArgSerSer 87
 |||||
 467 AGAAGATCTTCT 478
 |||||

seq_name: gb_est1:BB631162
 seq_documentation_block:

LOCUS
 DEFINITION
 BB631162 RIKEN full-length enriched, 16 days neonate thymus Mus
 musculus cDNA clone A130059E14 5', mRNA sequence.
 BB631162
 BB631162
 BB631162.1 GI:16467982
 EST.
 house mouse.
 Mus musculus
 Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 620)
 ARAKAWA, T., CARNINCI, P., FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A.,
 HIRAMOTO, K., HORI, F., ISHII, Y., ITO, M., KAWAI, J., KONNO, H., KODA
 M., KOYA, S., MATSUYAMA, T., MIYAZAKI, A., NOMURA, K., OHNO, M., OKAZAKI,
 Y., OKIDO, T., SAITO, K., SAKAI, C., SAKAI, K., SANO, H., SASAKI
 T., SHIBATA, K., SHINGAWA, A., SHIRAKI, T., SOGABE, Y., SUZUKI, H.,
 TAGAMI, M., TOGAWA, A., TAKEHISHI, F., TAKEEDA, Y., TANAKA, T., TOYA, T.,
 MURAMATSU, M. and HAYASHIZAKI, Y.
 RIKEN Mouse ESTs (ARAKAWA, T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yananaka, I., Alzawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.

FEATURES
 Location/Qualifiers
 1. 620
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="A130059E14"
 /clone_lib="RIKEN full-length enriched, 16 days neonate
 thymus"
 /tissue_type="thymus"
 /dev_stage="16 days neonate"
 /lab_host="DH10B"
 /note="Site 1: Sall; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse

117 oserProCysGlnAlaPheAsnHisTyrLeuSerAlaMetAlaSerM 134
 |||||||
 431 AAGTCTCTCTGCCAGCCCTTCAACCACTATCTCAGTGCATGCATCAGT 480
 |||||||
 134 etArgGlnAlaGluProAlaAspMetArgProGluIleTTPleAlaGln 150
 |||||||
 481 TGGAGATCTTACCAAA.....GTGGCACAA 506
 |||||||
 151 GluLeuArgArgIleGlyAspGluPheAsnAlaTyrTyr 163
 |||||||
 507 AATATCCAC.....GTGAT.....GCCTGGTAC 530
 |||||||

seq_name: gb_est1:AI971169

seq_documentation_block:
 LOCUS AI971169 492 bp mRNA linear EST 08-MAR-2000
 DEFINITION w24h12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2488679 3'
 similar to TR:043522 043522 BML. [1] ;, mRNA sequence.

ACCESSION AI971169
 VERSION AI971169
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 492)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Tumor Gene Index

COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 712 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 450.

FEATURES

Source 1..492

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2488679"

/clone_lib="NCI_CGAP_Pr28"

/sex="male"

/dev_stage="adult"

/lab_host="DH10B"

/note="Organ: prostate; Vector: pTT30-Pac (Pharmacia)

with a modified polylinker; Plasmid DNA from the

normalized library NCI_CGAP_Pr22 was prepared, and ss

circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs

985608-986759, 1101192-1101959, and 1217928-1220615).

Subtraction by Bento Soares and M. Fatima Bonaudo.

BASE COUNT 119 a 107 c 130 g 134 t 2 others

ORIGIN

alignment_scores:

Quality: 335.00 Length: 131

Ratio: 4.786 Caps: 1

Percent Similarity: 53.435

alignment_block:

US-09-508-832-10 x AI971169/rev

Align seg 1/1 to reverse of: AI971169 from: 1 to: 492

1 MetAlaLysGlnProSerAspValSerSerGluCysAspArgGluGlyAr 17
 |||||||
 425 ATGGCAAGCAACCTTCTGATGAAGTGTGAGTGCACCGAAGAGTAG 376
 |||||||
 17 gGlnLeuGlnProAlaGluArgProGlnLeuArgProGlyAlaProt 34
 |||||||
 375 ACAATTGCAGCTCGGAGAGGCTCCCAAGCTCAGACCTGGGGCCCTA 326
 |||||||
 34 hrSerLeuGlnThrGluProGlnGlyAsnProGluGlyAsnHisGlyGly 50
 |||||||
 325 CCTCCCTACACAGAGCCACAA..... 303
 |||||||
 51 GluGlyAspSerCysProHisGlySerProGlnGlyProLeuAlaProPr 67
 |||||||
 303 303
 |||||||
 67 oAlaSerProGlyProPheAlaThrArgSerProLeuPheIlePheMeta 84
 |||||||
 303 303
 |||||||
 84 rgArgSerSerLeuLeuSerArgSerSerGlyTyrPheSerPheasp 100
 |||||||
 303 303
 |||||||
 101 ThrAspArgSerProAlaProMetSerCysAspLysSerThrGlnThrPr 117
 |||||||
 302 ...GACAGAGCCAGCAGCCATGCTGTGCAAAATCAACAACACNC 256
 |||||||
 117 oSerProCysGlnAlaPheAsnHisTyrLeuSerAlaMet 131
 |||||||
 255 AAGTCTCTCTGCCAGGCCCTCAACCACTATCTCAGTGCATG 213
 |||||||

seq_name: gb_est1:AW629314

seq_documentation_block:

LOCUS AW629314 664 bp mRNA linear EST 31-MAR-2000

DEFINITION h156602.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone

IMAGE:2976314 3' similar to TR:043522 043522 BML. [1] ;, mRNA

sequence.

ACCESSION AW629314

VERSION AW629314.1 GI:7376104

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 664)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -400P from Gibco

High quality sequence stop: 458.

Location/Qualifiers

1..664

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2976314"

/clone_lib="Soares_NFL_T_GBC_S1"

/lab_host="DH10B"

/note="Organ: pooled; Vector: pTT30-Pac (Pharmacia) with

a modified polylinker; Site: 1: Not I; Site: 2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NBHL19W, testis NHT, and B-cell

NCI_CGAP_GCB1) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as

